UC Berkeley
UC Berkeley Electronic Theses and Dissertations

## Title

Achieving Plant Resistance to Pseudomonas syringae through Natural Diversity of Immune Receptors and Rational Design of ZAR1 Recognition Pathways

## Permalink

https://escholarship.org/uc/item/2066p0c1

## Author

Diplock, Nathan

## Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

# Achieving Plant Resistance to *Pseudomonas syringae* through Natural Diversity of Immune Receptors and Rational Design of ZAR1 Recognition Pathways

By

Nathan Diplock

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jennifer D. Lewis, Co-chair
Professor Yangnan Gu, Co-chair
Professor Patrick Shih

Spring 2024

Abstract

Achieving Plant Resistance to *Pseudomonas syringae* through Natural Diversity of Immune Receptors and Rational Design of ZAR1 Recognition Pathways

By

Nathan Diplock

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Jennifer D. Lewis, Co-chair

Professor Yangnan Gu, Co-chair

Plant pathogens cause significant crop losses, and current pathogen control methods are often detrimental to human and environmental health. Identifying pathogen resistance in cultivars or wild relatives of crops remains the primary approach for achieving pathogen resistance. However, mining natural resistance is time consuming, and population level resistance may not exist for emerging pathogens. As an alternative to traditional resistance breeding, natural molecular diversity can be used to understand the "rules" governing pathogen recognition, which can enable the design of novel resistance within existing recognition frameworks.

The intracellular angiosperm immune receptor, HOPZ-ACTIVATED RESISTANCE 1 (ZAR1), recognizes a remarkable diversity of pathogen secreted proteins, known as type 3 secreted effectors (T3SE). ZAR1 interacts with and monitors RLCKs for T3SE modification, including HOPZ ETI-DEFICIENT 1 (ZED1) and ZED1-RELATED KINASES (ZRKs). By monitoring diverse RLCKs, ZAR1 mediates broad immunodiversity through indirect effector interaction. Moreover, ZAR1 is the most highly conserved plant NLR described to date, suggesting the mechanism ZAR1 uses to recognize effectors is effective and warrants further investigation.

Chapter 1 introduces the rationale for my approach using natural protein diversity to understand the molecular mechanisms of plant pathogen resistance. In Chapter 2, I describe a natural diversity screen using *Arabidopsis thaliana* ecotypes to investigate the molecular mechanism of ZAR1 and ZED1-mediated recognition of the *Pseudomonas syringae* T3SE HopZ1a. This work demonstrates key regions in ZAR1 that are required for ZED1 interaction and HopZ1a recognition. In Chapter 3, I used natural protein diversity of *A. thaliana* RLCKs to investigate ZAR1 interaction specificity. I showed that single mutations in ZAR1 guardees can increase interaction strength with ZAR1, which is a crucial first step towards rational design of a ZAR1 guardee. Moreover, I showed that a *Nicotiana benthamiana* ZAR1 ortholog causes an autoactive HR phenotype in response to expression of certain *A. thaliana* ZRKs.

In Chapter 4, I generated a ZED1 point mutant library, and screened for HopZ1a recognition through *A. thaliana* ZAR1 and a *Nicotiana benthamiana* ZAR1 ortholog. While I

identified some functional divergence between ZAR1 orthologs, these ZAR1 orthologs are highly similar. I identified multiple mutations that abolish interspecies ZAR1-ZRK autoactivity. I also identified a mutation in ZED1 that abolishes autoactivity through tomato ZAR1, while still maintaining the ability to mediate HopZ1a recognition through tomato ZAR1. This advance is the first example of rational expansion of effector recognition through ZAR1 orthologs from non-model species.

In Chapter 5, I explored natural resistance to the highly virulent bacterial pathogen *P. syringae* pv. *tomato* race 1 (*Pst* race 1). Using a mapping population created by crossing a resistant wild tomato relative *Solanum pimpinellifolium* with a susceptible cultivated tomato variety, I used a novel sequencing approach to identify candidate genes contributing to *Pst* race 1 resistance. Traditional QTL mapping approaches assisted in narrowing down the list of candidate genes, for future evaluation in breeding programs.

In Chapter 6, I used single chain antibodies to attempt to achieve ZAR1 activation through antigen binding. Using two linker design approaches, I generated ZRK-GFP nanobody fusions with tailored conformations. I showed that ZRK-GFP nanobody fusions can successfully interact with ZAR1 and the GFP antigen. Although no ZRK-GFP nanobody fusions activate ZAR1 in the presence of the GFP antigen, future protein design tools may enable the success of this approach.

Overall, the work presented here elucidates key molecular mechanisms of ZAR1 mediated recognition of *P. syringae* effectors by leveraging natural diversity. It identifies candidate genes contributing to *P. syringae* resistance in a tomato mapping population. Finally, it explores approaches to novel ZAR1-mediated effector recognition.

# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS……II**

**CHAPTER I:** Introduction**……1**

**CHAPTER II:** A natural diversity screen in *Arabidopsis thaliana* reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex**……19**

**CHAPTER III:** Utilizing natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1**……52**

**CHAPTER IV:** Molecular dissection of the ZED1 pseudokinase enables expansion of HopZ1a recognition to *Solanum lycopersicum* ZAR1**……96**

**CHAPTER V:** *Solanum pimpinellifolium* exhibits complex genetic resistance to *Pseudomonas syringe pv. tomato***……132**

**CHAPTER VI:** Computational protein linker design for ZAR1-guardee nanobody fusions does not result in activation of plant immunity in the presence of antigen**……171**

**APPENDIX I:** *A. thaliana* Wall-Associated Kinase-Like proteins are targeted by HopZ1a**….190**

**APPENDIX II:** Generation of a *N. benthamiana ZAR1* knockout using CRISPR-Cas9**……...197**

## ACKNOWLEDGEMENTS

The work presented here is a masochistic passion project, made possible by the support of incredible scientists, peers, family and friends.

None of this research would have happened without the support of my advisor, Jennifer Lewis. When I joined the Lewis Lab as a rotation student, I had effectively no plant or molecular biology experience. Throughout the last five years, Jenn helped me become a better scientist, communicator, writer, mentor, mentee, and teacher. Jenn is an extremely patient mentor and cultivated a relationship and lab environment in which I felt comfortable asking for help and felt supported through both failures and successes. Jenn believed in me, and pushed me to do my best work, without suggesting I sacrifice my personal wellbeing. And lastly, Jenn was always willing to entertain my quasi-manic "ideas", none of which worked (see chapter 6). Thank you, Jenn. Additionally, thank you to Yangnan Gu and Patrick Shih, for your feedback which helped direct this dissertation.

The members of the Lewis Lab were instrumental in supporting me through this work. Thank you to Maël Baudin, for teaching me nearly all the lab techniques used in the following chapters. Thank you to Jana Hassan for being a massive support during the pandemic lockdowns, my confidant in complaining about the U.S political climate and state of the PGEC infrastructure, and a mentor always willing to share your years of practical tomato and *Arabidopsis* wisdom. Thank you to Marlo, for sharing your lab petri dish pets and your unfiltered views on pretty much everything. Thank you to Ilea, for all the support in navigating the grad student experience. And finally, thank you my undergraduate mentees; Helena Heiberger, Derek Xiang, Gabrielle Husain, and Rachelle Punsalan, for working with me on some pretty grueling and monotonous experiments, all while reminding me how exciting the research we do really is.

Thank you to my human, Sadie Mae Palmatier, for loving and supporting me unconditionally. You encourage me to advocate for myself and keep things in perspective. Going through this life with you is a gift and means more to me than I can put into words (though I will try to do so for our wedding vows).

Thank you to my mom and dad; neither of you knew why the sky is blue, or why the tide goes in and out, but you both encouraged me to explore my own theories. Figuring out the mysteries of the world around me is still what drives my curiosity. You both nurtured my curiosity, believed in me, and supported me as I followed my passion. Thank you to Jay Palmatier and Heidi Davis for making me a part of the family, always celebrating my accomplishments even when I felt they weren't worth celebrating, and your unwavering support throughout the past five years. Thank you to my brother Ryan and my sister Tesla. And finally, A huge thanks to all my friends who kept me sane through the pandemic, accompanied me on so many crazy adventures, and always encouraged me to take the morning off when the waves were firing.

Some odds and ends; thank you to the Sierra Nevada mountains, where I spent a combined half-year over the past five, for giving me the purest joy and calmest peace I have experienced in my adult life. To Craigslist, Costco, and Berkeley free piles, for allowing me to live a comfortable (even bougie?) life on a grad student stipend. And finally, thank you to the plants, especially

those with ZAR1, for always keeping me guessing and interested. You are intelligent beings in ways that I will never be able to comprehend. Keep your secrets!

# CHAPTER I: Introduction

Plant pathogens cause significant agricultural losses, threatening global food security[1]. Current pathogen control measures are often detrimental to human and environmental health. Engineering pathogen resistant crops is an appealing approach to address food insecurity and reduce the harm caused by agriculture. For crops with population level resistance in cultivars or wild relatives, resistance can be identified using natural diversity and introgressed through breeding or biotechnology[2]. However, emerging pathogen threats with no apparent population level resistance require novel approaches to achieving crop resistance[3]. Recognition of pathogens through intracellular immune receptors is an effective and durable pathogen recognition strategy.

## 1. Plant innate immunity

### *1.1 Overview of the plant immune system*

Plants only possess an innate immune system, which functions in pathogen recognition through two interconnected tiers. At the cell surface, pattern recognition receptors (PRRs) recognize pathogen-associated molecular patterns (PAMPs) such as flagellin, and mediate pattern-triggered immunity (PTI)[4]. The signal for PTI is transduced intracellularly through receptor-like cytoplasmic kinases (RLCKs), leading to intracellular $Ca^{2+}$ bursts, mitogen-activated protein kinase (MAPK) cascades, production of reactive oxygen species (ROS), transcriptional changes, and deposition of callose to strengthen the cell wall[5]. Some plant pathogens, including the bacterial pathogen *Pseudomonas syringae,* have a type III secretion system (T3SS) which penetrates the plant cell wall and membrane enabling secretion of type III secreted effectors (T3SE) into the plant cell[6–8]. These T3SEs typically promote virulence by suppressing PTI[4,8]. For example, the bacterial pathogen *Pseudomonas syringae* secretes numerous T3SEs, including AvrPphB which inhibits flagellin-triggered PTI by degrading an RLCK signaling component[9]. As a second tier of resistance, plants possess intracellular, nucleotide-binding leucine-rich repeat receptors (NLRs) that recognize T3SEs. When NLRs are activated in the presence of an effector, they mediate effector-triggered immunity (ETI) which is often accompanied by a programmed cell death known as the hypersensitive response (HR)[10]. The HR is effective at containing pathogen spread within a limited area of the plant[11,12].

Although PTI and ETI were once thought to function independently, recent research has shed light on crosstalk between these two defense strategies. For two separate *P. syringae* effectors, resistance through ETI in *Arabidopsis thaliana* was shown to occur only when PTI was also initiated[13,14]. While PTI appears to be required for full activation of ETI, the reverse correlation is also observed. PTI in response to a *P. syringae* flagellin PAMP was strengthened when ETI was induced by the T3SEs[15]. In addition, overexpression of an NLR-specific E3 ligase led to reduced accumulation of NLRs which compromised PTI[16]. Transcription of PTI components was also upregulated in plants following ETI[13,15]. Although ETI and PTI are often studied in isolation and depicted as separate responses, plant cells achieve the most effective immune response when both PTI and ETI are activated[17].

### *1.2 Plant NLR classifications, models of recognition, and signaling*

Plant NLRs are typically divided into three major categories depending on their N-terminal signaling domain. All canonical NLRs have a C-terminal leucine-rich repeat (LRR) domain, and a central Nucleotide-Binding, Apaf-1, R protein, Ced-4 (NBARC) domain. Canonical NLRs contain either a coiled coil (CC), Toll/Interleukin 1 receptor (TIR) or

Resistance to Powdery Mildew 8-like coiled coil ($CC_R$) at the N-terminus. Variation in CC-domain architecture is observed across species, including Solanaceae-domain (SD)-type CC domains, and Solanaceae I2-like CC domains[18]. Plant NLRs may also contain non-canonical "integrated domains" that function in effector recognition or NLR signalling[19]. Based on the N-terminal domain, NLRs are classified as CC-NLRs (CNLs), TIR-NLRs (TNLs) or $CC_R$-NLRs (RNLs). RNLs are often referred to as "helper" NLRs and are activated following pathogen recognition through a "sensor" TNL or CNL[20]. While CNLs and RNLs are present in all flowering plants, TNLs have been lost in monocots[21,22]. Although many NLRs function as sensor-helper pairs, some NLRs including ZAR1 and RESISTANCE TO *P. SYRINGAE* PV. MACULICOLA 1 (RPM1) mediate ETI independently of helper NLRs[23]. Upon activation, RNLs, TNLs and CNLs form higher order oligomeric structures leading to confirmational changes and activation of the N-terminal domain and downstream signaling.

To detect pathogens and trigger ETI, NLRs can monitor host proteins for T3SE modification, or interact directly with T3SEs. Direct NLR-T3SEs recognition usually occurs in a gene-for-gene pairing, where one NLR is responsible for recognition of a specific T3SE[24]. A unique type of direct interaction occurs through NLRs containing atypical integrated domains (NLR-IDs)[24]. The integrated domains in NLR-IDs usually mimic T3SE targets and are modified directly by T3SEs[25,26]. Upon modification, NLR-IDs typically mediate ETI through an additional NLR lacking IDs[25,26]. Although once described as non-canonical, T3SE recognition through NLR-IDs is widespread across plant species[25,26].

While the original model of plant immunity described gene-for-gene recognition as the primary method of T3SE recognition[27], recent research suggests this model is the exception rather than the rule for bacterial pathogens[28]. A far more common strategy is the monitoring of host proteins involved in PTI, or of mimics of PTI components through one or more NLRs[28]. This method of indirect T3SE detection is often referred to as the guardee model of resistance, where host proteins are "guarded" by NLRs. Guarded host proteins include PTI components, in addition to "decoy" proteins with no apparent function except as guardees[24]. As an example, receptor-like cytoplasmic kinases (RLCKs) are important signaling components for PTI[9], and common targets of pathogen effectors[29]. As such, several NLRs guard RLCKs for modification of pathogen effectors, including the *Arabidopsis thaliana* NLRs HOPZ-ACTIVATED RESISTANCE 1 (ZAR1), RPS5 and the *Solanum pimpinellifolium* NLR *PSEUDOMONAS* RESISTANCE AND FENTHION SENSITIVITY (Prf) [30–33]. Both direct and indirect recognition of T3SEs leads to complex downstream signaling.

Downstream ETI signaling is dependent on the class of NLR that is activated. Two well-studied TNLs, RESISTANCE TO *PERONOSPORA PARASITICA* 1 (RPP1) and RECOGNITION OF XOPQ1 (Roq1) form homo-tetramers upon effector recognition[34,35]. TIR domain association upon oligomerization produces an NADase active site, which results in the production of a variety of small molecules leading to activation of RNLs and further downstream signaling[35–39]. The well-studied CNLs, ZAR1 and STEM RESISTANCE 35 (Sr35) both form homo-pentamers upon activation[40–42]. CC domain association upon oligomerization in ZAR1 and Sr35 produces a membrane-associated non-selective cation channel, that promotes an influx of diverse cations including calcium[42,43]. Although no complete RNL structures have been described, $CC_R$ and CC domains share structural similarities[44]. A recent review provides an excellent description of plant NLR structural insights[45]. Like CNLs, RNLs form higher oligomeric structures upon activation, and activated RNLs function as membrane-associated cation channels[44]. While some NLRs such as ZAR1 and RPM1 function independently of RNLs,

they are also partially dependent on global TIR "sensor" signaling, suggesting NLR signaling may potentiate activation of other non-helper NLRs[11].

This introduction will focus on the use of natural protein diversity to explore the molecular mechanisms of plant recognition of effectors, recent progress towards the rational design of pathogen recognition, and rationales for future emphasis on engineering novel indirect pathogen recognition. I will emphasize examples of indirect pathogen recognition, as these recognition systems offer advantages in the context of crop breeding for bacterial pathogen resistance. I will explore natural protein diversity in the context of indirect recognition at three different scales 1) natural diversity across species 2) natural diversity within species, and 3) natural diversity within individuals. Finally, I will extent these insights to approaches for rational engineering of plant-pathogen recognition.

## 2. Natural diversity to understand molecular mechanisms of indirect recognition

### *2.1 Natural diversity across species*

Plants typically possess hundreds of NLRs, with high diversity between and within species[21,22,46]. High diversity of plant NLRs enables broad immunodiversity between closely related species and within a species at the population level. By mining natural NLR diversity, pathogen resistance can be identified and introgressed into other closely related populations. A classic example of successful identification and introgression of NLR-based resistance is the use of the *Solanum pimpinellifolium* pathogen resistance *Pto*/*Prf* cluster in cultivated tomato[47]. The *Pto/Prf* cluster was introgressed into *Solanum lycopersicum* in the 1930s and provided successful resistance to *P. syringae* race 0 strains. Two *P. syringae* effectors, AvrPto and AvrPtoB from *P. syringae* race 0 strains target PTI components to suppress plant resistance[48–50]. Two RLCKs, Pto and Fen mimic PTI components targeted by AvrPto *and* AvrPtoB[51]. Within the *Pto/Prf* cluster, the NLR Prf guards Pto and Fen, leading to recognition of AvrPto and AvrPtoB[32,52,53]. The efficacy of the *Pto/Prf* cluster in cultivated tomato demonstrates the value of introgressing resistance from diverse wild relatives. While natural diversity within the *Pto/Prf* cluster exists[54], it has not yet been utilized to investigate mechanisms underlying indirect recognition, compared to the examples that follow.

NLR diversity in addition to high copy number variation makes identifying NLR orthologs across plant species challenging. In addition, the transfer of NLRs form one species to another can result in autoactivation of plant immunity, or unsuccessful transfer of resistance from the original host[55–59]. Host factors that function in indirect recognition of pathogen effectors follow the same trend as NLRs, with variable conservation across species[60,61]. Expressing NLRs that function through indirect recognition in heterologous systems can lead to autoactivation of the HR due to lack of guardee suppression[55,56,58]. For example, the *A. thaliana* NLR, RPS2 is activated when its guardee RPM1-INTERACTING 4 (RIN4) is degraded by pathogen effectors[55,62,63]. When RPS2 is transiently expressed in *N. benthamiana*, it results in an autoactive HR since no RIN4 orthologs exist in *N. benthamiana* to keep RPS2 in an off-state[55]. While NLR and guardee variability and diversity can be detrimental in certain contexts, it can also be used as a tool to explore the molecular mechanisms of NLR indirect recognition.

Most NLRs are described in one species and are not present or functionally conserved outside of that species. The highly conserved angiosperm NLR, ZAR1, is an exception to typical pattern of NLR variability across species. ZAR1 orthologs exist in nearly all eudicots[61,64–66]. Moreover, ZAR1 orthologs are functionally conserved, meaning combinations of ZAR1 orthologs and guarded RLCKs can function across species[30,67]. In *A. thaliana*, ZAR1 guards the

RLCK, HOPZ ETI-DEFICIENT 1 (ZED1), which triggers ETI when modified by the *P. syringae* effector HopZ1a[30,64]. When ZED1 is transiently expressed in *N. benthamiana,* it can mediate HopZ1a recognition through an *N. benthamiana* ZAR1 ortholog (NbZAR1)[67]. The functional conservation of ZAR1 and the high diversity of ZED1-RELATED KINASE (ZRK) guardees[61,68–70] make the ZAR1/ZRK1 module ideal for understanding the molecular mechanisms of pathogen recognition, and an appealing system for rational design of effector recognition[71]. Recently, nine ZAR1 orthologs were shown to have variable autoactive HR phenotypes when transiently expressed with ZRKs in *A. thaliana*[61]. Furthermore, two out of thirteen *A. thaliana* ZRKs have been shown to result in *N. benthamiana* ZAR1 autoactivation when transiently expressed in *N. benthamiana*, demonstrating that autoactivation is likely dependent on interspecies combinations of ZRKs and ZAR1. In the future, variation in autoactivation resulting from natural ZRK diversity may provide insight into the mechanisms of ZAR1 autoactivation (see Chapter 3)[61,71]. Recently, a ZED1 mutant library was used to explore functional divergence between ZAR1 orthologs, the molecular mechanism of autoactivity, and the feasibility of maintaining effector recognition while abolishing autoactivity (see Chapter 4). Mutations in ZED1 that perturbed AtZAR1 recognition of HopZ1a generally also perturbed NbZAR1 recognition of HopZ1a. However, two ZED1 mutations resulted in a lack of HopZ1a recognition through AtZAR1 but not NbZAR1. Further analysis suggested AtZAR1 requires more residues of ZED1 to be acetylated by HopZ1a to initiate activation when compared to NbZAR1. Furthermore, a characterization of ZED1 mutants in the context of autoactivation of NbZAR1 suggested that ZAR1-ZRK autoactivity likely results from the ZRK activation loop. A mutation that abolished NbZAR1 autoactivity but not HopZ1a recognition was used to expand HopZ1a recognition to a ZAR1 ortholog in tomato (see Chapter 4). Overall, ZAR1 remains a unique system for using natural diversity to understand the molecular mechanisms of a key pathogen recognition hub and generate novel interspecies resistance combinations. A recent in-depth review of ZAR1 pathogen recognition highlights the expansive ZAR1 recognition network in detail[72].

While ZAR1 and ZAR1-guarded RLCKs are highly conserved, other NLRs that recognize pathogens indirectly through host factors are divergent or absent across species. The *A. thaliana* NLR, RPS2, recognizes AvrRpt2 through the degradation of the guarded host factor RIN4[55,62,63]. Neither RPS2 or RIN4 are present in *N. benthamiana*[55]. When RPS2 is transiently expressed in *N. benthamiana*, it results in an autoactive HR, due to the absence of RIN4-mediated suppression of RPS2[55]. While RIN4 functions in AvrRpt2 recognition through RPS2 in *A. thaliana*, RIN4 is also required for AvrRpm1 recognition through RPM1[73]. Although RIN4 is absent in *N. benthamiana*, four RIN4-like proteins are found in soybean (*Glycine max*) where they mediate indirect recognition of the *P. syringae* AvrRpm1 through unrelated NLR, RPG1-B[74,75]. One of these RIN4-like proteins, GmRIN4b is able to complement the *A. thaliana rin4* mutation and mediate recognition of *P. syringae* AvrRpm1, while a different RIN4-like protein, GmRIN4a is not[74]. Similarly, an additional RIN4 ortholog was identified in apple (*Malus domestica*), that mediates indirect recognition of *Erwinia amylovora* effector AvrRpt2 through another independently evolved NLR, MR5[76]. By comparing AtRIN4 with MdRIN4, the authors were able to identify two residues, AtRIN4$^{N158}$ and AtRIN4$^{Y163}$, that are crucial for interaction and regulation of RPS2, RPM1 and MR5[76]. Using domain swaps, the authors identified regions of MdRIN4 and AtRIN4 that confer recognition specificity for AvrRpt2 from *P. syringae* and *E. amylovora*. Additional comparisons of 266 RIN4 orthologs across plant species, revealed that RIN4 orthologs can largely be grouped by polymorphisms at two key RIN4 regulatory sites

N158 and Y163, termed the RIN4 specificity motif (RSM)[60]. Functional characterization of 33 RIN4 orthologs selected from diverse species demonstrated that the RIN4 RSM underlies RIN4-NLR compatibilities[60]. Finally, functional characterization of RIN4 orthologs was used to generate RIN4 chimeras that broadened RIN4 NLR regulation specificity[60].

Natural diversity of the RESISTANCE TO *P. SYRINGAE* 5 (RPS5) guarded protein AVRPPHB SUSCEPTIBLE 1 (PBS1) has aided our understanding of the molecular mechanisms of an NLR-guarded PTI component. The *P. syringae* effector AvrPphB cleaves the RLCK, PBS1 that triggers RPS5-mediated immunity[31]. While PBS1 is highly conserved across plant species[77], likely due to a function in PTI[9,78], the NLRs that guard PBS1 across plant species are not orthologs[79,80], similar to the NLRs guarding RIN4 across species (see above). The PBS1 SEMPH motif is required for RPS5 activation upon AvrPphB cleavage. The identification of the PBS1 SEMPH motif using natural protein diversity within an individual is discussed in a later section (see *Natural diversity within individuals*). To further explore the importance of the PBS1 SEMPH motif, Sun and colleagues analyzed proteins related to wheat (*Triticum aestivum*) PBS1 (TaPBS1)[81]. Unlike the AtPBS1 SEMPH, TaPBS1 has a STRPH motif, which is conserved across other monocot PBS1 homologs[81]. While the TaPBS1 STRPH motif does not affect AvrPphB cleavage or interaction with *A. thaliana* RPS5, it does inhibit TaPBS1 activation of *A. thaliana* RPS5 upon AvrPphB cleavage[81]. Replacing the TaPBS1 STRPH motif with the AtPBS1 SEMPH motif allows for functional AvrPphB recognition through RPS5[81]. In addition to AvrPphB recognition, PBS1 homologs from wheat and barley have been used to explore *A. thaliana* RPS5 substrate specificity[80,81]. Understanding PBS1 molecular mechanisms led to the earliest example of successful rational design of pathogen recognition (see *Toward rational design of effector recognition*).

*2.2 Natural diversity within species: A. thaliana ecotypes*

Utilizing natural diversity to identify resistance within a species at the population level is a widespread practice in plant resistance breeding, and this topic has been reviewed previously[2,82,83]. Here, we will focus on examples where natural ecotype diversity has been used to understand the molecular mechanisms of *A. thaliana* immunity. The *A. thaliana* 1001 genomes project, in addition to numerous recombinant inbred lines (RILs) resulting from ecotype crosses[84], are an unparalleled resource for connecting a species pan-genome with population level variation in phenotypes[85].

When variation in pathogen recognition exists across *A. thaliana* ecotypes, this variation can be used to conduct genome-wide association studies (GWAS) to pinpoint mechanisms of pathogen recognition[86]. Auon and colleagues used a genome wide association (GWA) approach to identify new mechanisms of heat tolerant resistance to *Ralstonia solanacearum* in *A. thaliana*[87]. By screening 176 natural accessions of *A. thaliana* for resistance to *R. solanacearum* at high temperature, two loci were identified. One loci contained a previously described NLRs RPS4/RRS1-R, and an additional loci containing a STRICTOSIDINE SYNTHASES-LIKE 4 (SSL4) protein[87]. The requirement of SSL4 was validated further in the context of heat tolerant *R. solanacearum* resistance, demonstrating a novel class of proteins involved in plant pathogen resistance, although the mechanisms of this resistance is still poorly understood[87].

Similar to the approach used to identify heat tolerant *R. solanacearum* resistance, Huard-Chauveau and colleagues used recombinant inbred lines (RILs) to identify a novel component in *Xanthomonas campestris* resistance[88]. *A. thaliana* Col-5 and Kas-1 ecotypes exhibit contrasting resistance phenotypes upon *X. campestris* infection[88]. Using a Col-5 x Kas-1 RIL population[84], in

addition to a GWA approach with 381 *A. thaliana* accessions, the authors identified a RLCK in the ZRK family, *RESISTANCE RELATED KINASE 1* (*RKS1*)[88]. Further exploration of *RKS1* sequences across 95 *A. thaliana* ecotypes identified signatures of balancing selection, suggesting *RKS1* is important in pathogen recognition[88]. Further research described the role of RKS1 in recognizing the *X. campestris* effector AvrAC through PBS1-LIKE 2 (PBL2), resulting in ETI-mediated by ZAR1[89].

While *A. thaliana* ecotypes can be used to identify novel mechanisms of resistance associated with a specific pathogen, natural protein diversity across ecotypes can also be used to explore the molecular mechanisms of a known resistance system. To explore the molecular mechanisms of *P. syringae* recognition mediated by ZAR1, Baudin and colleagues screened nearly 300 *A. thaliana* ecotypes for the HR in response to *P. syringae* carrying *hopZ1a*[90]. Since previous research has demonstrated HopZ1a is indirectly recognized by ZAR1 via ZED1[30,64], the authors were able to leverage data from the 1001 genomes project to prioritize testing ecotypes with non-redundant ZAR1 and ZED1 alleles. Further characterization of *A. thaliana* alleles unable to recognize HopZ1a identified residues essential for ZAR1-ZED1 interaction and ZAR1 function[90]. This will be discussed further in Chapter 2.

While crosses between certain *A. thaliana* ecotypes are compatible, crosses between other ecotypes can result in hybrid incompatibility which often results in an autoactive immune signaling phenotype[91]. Hybrid incompatibility is observed across plant species, and has been reviewed extensively[92,93]. By intercrossing 80 *A. thaliana* ecotypes and mapping, Barragan and colleagues were able to identify alleles of the NLR RESISTANCE TO *PERONOSPORA PARASITICA* 7 (RPP7) and a non-NLR resistance protein RESISTANCE TO POWDERY MILDEW 8 (RPW8) which are incompatible[91]. By swapping domains from different *A. thaliana* RPW8 alleles, the authors demonstrated that the difference in lengths of the RPW8 C-terminal repeats contribute to RPP7-mediated hybrid necrosis[91].

In addition to identifying specific QTL associated with resistance to pathogens, the mechanisms of NLR activation, and NLR-mediated hybrid necrosis, natural diversity across populations can be used to study broader models of plant resistance, and trade-offs between fitness and immunity. *Cucumber mosaic virus* is a common pathogen of *A. thaliana* in natural populations. Montes and colleagues explored co-evolution of resistance and virulence in natural populations of *A. thaliana* using different strains of *Cucumber mosaic virus*[94]. While host-pathogen co-evolution is often assumed in the interaction between biotrophic pathogens and plants, the *A. thaliana* and *Cucumber mosaic virus* relationship is a rare example of experimental data demonstrating that a virus exerts selection pressure on natural plant populations as a result of fitness detriments from infection[94]. Moreover, resistance to *Cucumber mosaic virus* was not correlated with climatic or life history characteristics, suggesting that resistance in this system is independent of abiotic and life history variation[94]. To further examine the association between life history strategies and immunity outside of the *Cucumber mosaic virus* system, Glander and colleagues explored gene expression across ecotypes of *A. thaliana*[95]. Comparisons showed that alleles associated with a decrease in expression of immune components were correlated with shorter vegetative lifespans, suggesting that the *Cucumber mosaic virus* model may be an exception to a broader correlation between plant life history and resistance[95].

To investigate the fitness costs and benefits specifically associated with NLRs, MacQueen and colleagues used *A. thaliana* ecotypes carrying *RPS2* alleles that confer resistance, or susceptibility to *P. syringae* carrying *avrRpt2*[96]. Both *RPS5* and *RPM1* show presence/absence variation across *A. thaliana* ecotypes, with susceptible ecotypes typically having deletions of

these NLRs[97,98]. Individuals carrying resistant *RPS5* or *RPM1* alleles incur a 5-10% decrease in fitness relative to individuals with susceptible alleles in the absence of pathogen infection[97,98]. Unlike *RPS5* and *RPM1*, *RPS2* exists in all *A. thaliana* ecotypes with either resistant or susceptible alleles[96]. MacQueen and colleagues demonstrated that no significant fitness differences exist between individuals carrying *RPS2* alleles associated with susceptibility or resistance[96]. Moreover, knocking out *RPS2* results in a large fitness detriment likely explained by its role as a negative regulator of defense gene expression and contrary to the *RPS5* and *RPM1* findings[96]. Taken together, natural diversity across *A. thaliana* demonstrates that NLR genomic architecture across populations can provide insights into fitness costs associated with resistance[96].

Overall, natural diversity within a species population can be used to identify genes associated with resistance[84] or autoimmunity[87] using GWAS. Natural diversity of specific alleles can be used to explore the molecular mechanism of pathogen recognition[90]. Using natural diversity within species to establish broader models of plant-pathogen interactions highlights the evolutionary complexities of these relationships and alludes to the complexity of recognition at the molecular level. While possessing certain NLRs may cause fitness detriments[97,98], other NLRs may be necessary for negatively regulating immunity and preventing fitness detriments[96]. While variation in life history may be unrelated to resistance for certain plant-pathogen interactions[94], this may not hold for all plant-pathogen interactions[95].

*2.3 Natural diversity within individuals*

While natural diversity of protein homologs can be explored across individuals within a population, or across species, natural protein diversity also exists within an individual across homologs that comprise a protein family. In plants, protein diversity is especially high within the RLCK[99] and NLR[21,22,46] protein families, and this natural protein diversity has been used to explore the molecular mechanisms of pathogen recognition.

The flax (*Linum usitatissimum*) *L* genes provide an excellent early example of utilizing protein diversity within individuals to identify mechanisms of gene-for-gene pathogen resistance. The flax *L* locus contains at least twelve closely related NLR alleles which confer resistance to flax rust (*Melampsora lini*)[100]. The NLRs in the *L* locus recognize specific *M. lini* avirulence effectors in a gene-for-gene manner[101]. By performing domain swaps between different *L* alleles, it was determined that amino acid variation in the LRR domains of *L* alleles was responsible for *M. lini* effector recognition specificity[100]. Additional interaction data validated that L proteins and *M. lini* effectors interact as highly specific pairs, dependent on the LRR domain[101].

While direct recognition of effectors is a common strategy for fungal pathogens, indirect recognition is the predominant strategy for detecting bacterial pathogens[28]. Natural diversity in the ZRK and PBL family of RLCKs has been used to explore indirect bacterial effector recognition through ZAR1. By guarding nearly the entire ZRK family in the absence of pathogen effectors, ZAR1 conveys incredible immunodiversity[71,102]. In a recent study characterizing the *P. syrinage* effector pan-genome, ZAR1 was found to convey resistance to around half of the 494 *P. syringae* strains characterized[28]. Each *A. thaliana* ZRK recognizes specific T3SE(s), which have diverse enzymatic functions, either directly or indirectly through PBL kinases[68,72,89]. The *A. thaliana* ZRK, RKS1, interacts directly with ZAR1 in the absence of pathogen effectors. Uridylylation of PBL2 by the *Xanthomonas campestris* T3SE AvrAC leads to a PBL2-RKS1 interaction, and activation of ZAR1[89]. In contrast, ZED1 is directly acetylated by the *P. syringae* T3SE HopZ1a[68], although two PBLs, SZE1 and SZE2, have been described as partially required

for ZAR1-mediated recognition of HopZ1a[103]. While the PBL interaction profile of ZED1 changes significantly upon activation, the functional relevance of this diverse PBL interaction profile is unclear[104]. While certain PBLs have known roles in PTI[9], ZRKs appear to be decoys with no function outside of ZAR1-mediated ETI[68,88]. Recent phylogenetic analysis of RLCKs revealed that ZRKs likely evolved from the intracellular kinase domains of wall-associated kinase-like (WAKL) proteins[61]. While the role of WAKLs in PTI has been described extensively across plant species[105], no WAKLs are known targets of T3SEs that are recognized through ZAR1. In contrast, AvrAC has been shown to extensively target PBLs, including those with a role in PTI[106], and is recognized by ZAR1 through PBL2. Comparative analysis of PBL2 and PBL3 revealed differences in two amino acids near the AvrAC acetylation site, which are required for PBL2 function[61]. By mutating these two amino acids in PBL3 to match the corresponding residues PBL2, Gong and colleagues were able to achieve AvrAC recognition through PBL3[61].

Recently, the natural diversity of *A. thaliana* RLCKs was used to demonstrate that ZAR1 interaction specificity extends beyond the ZRK family (see Chapter 3)[71]. Through transient expression in *N. benthamiana*, *A. thaliana* ZAR1 was shown to interact with nearly all ZRKs, the kinase domain of WAKL2, and PBL2, but not the more distantly related RLCK, BIK1[71]. Through protein truncations, the ZAR1 LRR domain was shown to be the determinant of ZAR1 interaction specificity across RLCKs[71]. By comparing the amino acid sequences of ZRKs, the authors were able to identify a single mutation in ZRK10 that strengthened interaction with ZAR1[71]. Plant kinases are common targets of T3SEs[29], and taken together, ZAR1 is highly effective at monitoring kinases that are PTI components (through PBLs) or mimics of PTI components (ZRKs)[72]. Demonstrating that mutations in ZAR1 guardees can lead to strengthened ZAR1 interactions[71] is a key proof of concept towards rationally designing a ZAR1 guardee from a T3SE kinase target (see Chapter 3).

In addition to its role in ZAR1 recognition, the PBL protein family has been used extensively to explore RPS5 interaction specificity, and the molecular mechanisms of AvrPphB recognition. PBS1 is a RLCK VII family protein, which in *A. thaliana* consists of 45 members, 29 of which are PBS1-like (PBL) proteins[9,99]. By comparing natural diversity of *A. thaliana* PBLs, Zhang and colleagues were able to validate the importance of the PBS1 GDK motif, which is required for PBS1 cleavage by AvrPphB and initiation of ETI through RPS5[9,31]. While nine tested PBLs containing the GDK motif were cleaved by AvrPphB, PBL6 was not cleaved, as it contains a D-to-S substitution in the GDK motif[9]. However, PBL1 which contains a D-to-E substitution in the GDK motif was cleaved normally, suggesting acidic residues are compatible[9]. Furthermore, the authors showed that a RLCK XII member, BSK1, is not cleaved by AvrPphB, demonstrating that AvrPphB has substrate specificity for PBLs (RLCK VII)[9]. To further determine AvrPphB's specificity for PBLs, Qi and colleagues tested the five PBLs most similar to PBS1 and the most dissimilar PBL to PBS1 for cleavage by AvrPphB. The authors found that only PBL27 produced a consistent cleavage product in the presence of AvrPphB[107]. Using natural diversity of PBLs, Qi and colleagues generated PBS1-PBL27 chimeras to explore RPS5 interaction specificity for PBS1. The chimeras demonstrated that the C-terminus of PBS1 was essential for RPS5 activation. Further amino acid comparisons identified a five amino acid motif unique to PBS1, termed the SEMPH motif, which was required for RPS5-mediated recognition of PBS1 cleavage products[107].

**3. Toward rational design of effector recognition**

Indirect recognition of pathogen effectors is a common strategy for NLR-mediated plant immunity. While the original gene-for-gene model of NLR-effector recognition still holds true for some effector recognition systems, indirect recognition appears to be a far more common NLR recognition strategy, particularly for bacterial pathogens[28]. By guarding PTI components, or mimics of PTI components, a single NLR can convey durable, broad immunodiversity. Therefore, rational design of indirect effector recognition may be a successful approach for durable broad-spectrum novel pathogen resistance. While examples of rational design of indirect recognition exist, rational design of direct NLR-effector recognition has been more common. This is likely due to decreased complexity of designing one interaction interface (NLR-effector) when compared to the two interfaces required for indirect recognition (NLR-guardee-effector). In this section, rational design of both direct and indirect T3SE recognition will be discussed, with an emphasis on design informed by natural protein diversity.

Modification of the PBS1 protease cleavage site has led to novel viral resistance in several plant species. By replacing the *A. thaliana* PBS1 AvrPphB cleavage domain with the recognition sequence of the tobacco etch virus (TEV) NIa protease, Kim and colleagues were able to achieve TEV-mediated cleavage of PBS1 and subsequent immune activation through transient expression in *N. benthamiana*[108]. Similarly, *A. thaliana* expressing PBS1 with the turnip mosaic virus (TuMV) NIa cleavage site resulted in TuMV resistance through RPS5[108]. Viral resistance in soybean and potato was successfully designed by modifying the PBS1 protease cleavage site and because PBS1 is conserved outside of model species. Rational design of SMV resistance in soybean was achieved using a soybean PBS1 ortholog with the AvrPphB cleavage site replaced with a soybean mosaic virus (SMV) NIa protease cleavage site[109,110]. Using the same approach, Bai and colleagues modified a potato PBS1 ortholog to contain the potato virus Y (PVY) NIa protease cleavage site, which mediated effective PVY resistance in potato[79]. These successful rational design approaches relied heavily on in-depth mechanistic understanding of PBS1 largely determined using natural protein diversity.

Rational design of direct effector recognition has been achieved successfully through modification of NLR LRR domain residues. The wheat NLR Sr50 directly binds and recognizes the wheat stem rust (*Puccinia graminis* f. sp. *tritici*) effector AvrSr50[111]. The closely related wheat NLR Sr33 shares nearly 80% sequence identity with Sr50, but cannot recognize AvrSr50[112]. By analyzing sequence variability in a structural context between Sr50, Sr33, and related homologs across Poeae, Brachypodieae and Triticeae, Tamborski and colleagues identified variable residues in the Sr33 LRR domain. Modifying 12 variable residues in Sr33 to match residues in Sr50 was sufficient to achieve AvrSr50 recognition through Sr33[112]. Using a similar approach informed by structural data, Förderer and colleagues identified key differences between the wheat NLR Sr35 and a barley (*Hordeum vulgare*) homolog SH1[42]. Sr35 directly interacts with AvrSr35, triggering ETI in wheat[113]. By swapping 10 residues in the Sr35-AvrSr35 binding interface into barley SH1, the resulting gain-of-function barley SH1 allele was capable of recognizing AvrSr35[42]. Engineered SH1 and Sr33 represent successful rational NLR engineering proof-of-concepts, which relied heavily on natural diversity in LRR domains.

While direct effector recognition through rationally designed LRR domains is a feasible approach for novel gene-for-gene resistance, targeted modification of NLRs with integrated domains such as heavy metal-associated (HMA) domains have achieved multi-effector resistance. The rice sensor NLRs RGA5 and Pikp-1 contain integrated HMA domains, which interact directly with sequence unrelated structurally similar effectors from *Magnaporthe oryzae* (rice blast fungus) [114–117]. RGA5 interacts with and recognizes AVR1-CO39 and AVR-Pia while

Pikp-1 recognizes AVR-PikD[114–116]. Importantly, AVR1-CO39 and AVR-Pia target a different surface of the RGA5$^{HMA}$ compared to the Pikp-1$^{HMA}$ surface targeted by AVR-PikD[118,119]. By comparing the AVR1-CO39 amino acid sequence with that of a closely related effector AvrPib, Liu and colleagues were able to identify two amino acids in the RGA5$^{HMA}$ that changed effector binding specificity from AVR1-CO39 to AvrPib[120]. In another study using a structure-informed engineering approach, residues in the Pikp1$^{HMA}$-AVR-PikD binding interface were integrated into the RGA5$^{HMA}$ domain, resulting in a novel RGA5 variant capable of recognizing Avr-PikD without disrupting RGA5-mediated AVR1-CO39 and AVR-Pia recognition[121]. However, extending RGA5-mediated disease resistance to *M. oryzae* carrying *AVR-PikD* in transgenic rice required variants with additional mutations outside of the RGA5$^{HMA}$ to ensure proper ETI activation[122]. Modifications to the Pikp-1$^{HMA}$ have focused on expanding recognition to novel effectors outside of natural Pikp-1 and RGA5 recognition repertoires. Through structure-informed mutagenesis, Pikp-1 recognition was expanded to alternative AVR-Pik alleles, including AVR-PikA AVR-PikC, AVR-PikE, and AVR-PikF[123,124].

While expanding NLR recognition specificities to closely related effectors may be an effective strategy for resistance to certain pathogens, the ability to tailor NLR recognition to any effector is a critical breakthrough. Variable domain heavy-chain-only antibodies, also known as nanobodies, are small proteins (~15 kDa) within the *Camelidae* species adaptive immune system. Due to their small size, approximately one-tenth the size of a traditional antibody, nanobodies have been used for diverse applications across biology[125]. Recently, nanobodies raised against either GFP or mCherry were swapped with the Pik-1$^{HMA}$ domain, resulting in Pik-1 recognition of GFP and mCherry[126]. Moreover, Pik-1 nanobody fusions conferred resistance to *Potato virus X* expressing either GFP or mCherry when transiently expressed in *N. benthamiana*[126]. Since nanobodies can be raised against virtually any target, this approach could be used to rapidly generate novel resistance to any pathogen effector through NLRs with similar domain architecture to Pik-1[126]. Similar approaches could be used to create nanobody-mediated immunity across diverse plant species.

## 4. Conclusions

While natural diversity has elucidated the molecular mechanisms of several pathogen recognition systems, many avenues exist for future research and questions and challenges remain. Understanding the relationship between the HR and resistance is crucial for successful rational design of pathogen recognition. The HR is often used as the phenotype used to screen for gain of recognition through novel alleles. However, rational design attempts have resulted in NLRs capable of mediating the HR, but not resistance in their host system[121]. Another outstanding question is why indirect recognition is the dominant strategy for plant recognition of bacterial pathogens[28].

Further investigation of the ZAR1 recognition system may answer both of these outstanding questions. While natural diversity in the ZAR1 recognition system has been used to identify loss and gain of function ZRK and PBL mutations, as well as loss of function ZAR1 mutations, there is great potential to use natural diversity of ZAR1, ZRK, and PBL proteins to investigate the molecular mechanisms of ZAR1 and ETI more broadly. For example, ZAR1 can mediate pathogen resistance with and without the HR through different ZRKs[72]. This could be used to explore the relationship between the HR and pathogen resistance, which is still poorly characterized. Moreover, diverse ZRKs and PBLs in *A. thaliana* and *N. benthamiana* recognize diverse bacterial effectors, and ZRKs from these species can function with other ZAR1

orthologs. It is likely that ZRKs from other species could expand the effector recognition repertoire outside of those currently known, for tailored resistance. Continuing to explore the molecular mechanisms associated with plant-pathogen recognition will advance efforts to rationally design pathogen recognition.

Changes in land use and climate threaten natural diversity. While numerous utilitarian and ethical arguments for conserving biodiversity exist, the decades of research highlighted here demonstrate that natural diversity is critical for achieving pathogen resistance in agriculture. Although novel mechanisms for achieving plant resistance to pathogens are rapidly evolving, they rely heavily on natural protein diversity. Leveraging plant natural diversity to generate crop pathogen resistance may reduce the harm caused by current plant-pathogen control approaches, which will have direct positive feedback on maintaining biodiversity.

In the following chapters, I describe my work leveraging natural diversity within individuals, populations, and across species to investigate the molecular mechanism of plant resistance to *P. syringae*. I also include my progress towards rational design of *P. syringae* recognition through ZAR1. In Chapter 2, I describe a natural diversity screen in *A. thaliana* ecotypes which identifies novel loss-of-function ZAR1 alleles. In Chapter 3, I use the natural diversity within *A. thaliana* RLCKs to describe ZAR1 interaction specificity. Here, I demonstrate that single amino acid substitutions in ZAR1 guardees can strengthen ZAR1-gaurdee interactions. In Chapter 4, I screen a ZED1 mutant library against three ZAR1 orthologs and show that ZAR1-ZRK interspecies autoactivity can be overcome through mutations in guardees without impacting ZRK effector recognition. In Chapter 5, I describe efforts to identify host resistance to the highly virulent *P. syringae* pv. *tomato* race 1 in an accession of the wild tomato species *Solanum pimpinellifolium*. Finally, in Chapter 6, I present an unsuccessful effort to integrate nanobodies into the ZAR1-RKS1 recognition system. Overall, this work contributes to our understanding of the molecular mechanisms of *P. syrinage* recognition and presents promising directions toward the rational design of effector recognition.

## Bibliography


1. Savary, S. *et al.* The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* **3**, 430–439 (2019).
2. Nelson, R., Wiesner-Hanks, T., Wisser, R. & Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **19**, 21–33 (2018).
3. Ristaino, J. B. *et al.* The persistent threat of emerging plant disease pandemics to global food security. *Proc. Natl. Acad. Sci. U. S. A.* **118**, 1–9 (2021).
4. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
5. Couto, D. & Zipfel, C. Regulation of pattern recognition receptor signalling in plants. *Nat. Rev. Immunol.* **16**, 537–552 (2016).
6. Tang, X., Xiao, Y. & Zhou, J. M. Regulation of the type III secretion system in phytopathogenic bacteria. *Mol. Plant-Microbe Interact.* **19**, 1159–1166 (2006).
7. He, S. Y., Nomura, K. & Whittam, T. S. Type III protein secretion mechanism in mammalian and plant pathogens. *Biochim. Biophys. Acta - Mol. Cell Res.* **1694**, 181–206 (2004).
8. Alfano, J. R. & Collmer, A. TYPE III SECRETION SYSTEM EFFECTOR PROTEINS: Double Agents in Bacterial Disease and Plant Defense. *Annu. Rev. Phytopathol.* **42**, 385–414 (2004).
9. Zhang, J. *et al.* Receptor-like cytoplasmic kinases integrate signaling from multiple plant

immune receptors and are targeted by a Pseudomonas syringae effector. *Cell Host Microbe* **7**, 290–301 (2010).
10. Heath, M. C. Hypersensitive response-related death. *Plant Mol. Biol.* **44**, 321–334 (2000).
11. Jacob, P. *et al.* Broader functions of TIR domains in Arabidopsis immunity. *Proc. Natl. Acad. Sci.* **120**, 2017 (2023).
12. Jacob, P., Hige, J. & Dangl, J. L. Is localized acquired resistance the mechanism for effector-triggered disease resistance in plants? *Nat. Plants* **9**, 1184–1190 (2023).
13. Yuan, M. *et al.* Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature* **592**, 105–109 (2021).
14. Ngou, B. P. M., Ahn, H.-K., Ding, P. & Jones, J. D. G. Mutual Potentiation of Plant Immunity by Cell-surface and Intracellular Receptors. *bioRxiv* (2020) doi:10.1101/2020.04.10.034173.
15. Ngou, B. P. M., Ahn, H. K., Ding, P. & Jones, J. D. G. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature* **592**, 110–115 (2021).
16. Tian, H. *et al.* Activation of TIR signalling boosts pattern-triggered immunity. *Nature* **598**, 500–503 (2021).
17. Yuan, M., Ngou, B. P. M., Ding, P. & Xin, X. F. PTI-ETI crosstalk: an integrative view of plant immunity. *Curr. Opin. Plant Biol.* **62**, 102030 (2021).
18. Wang, J., Han, M. & Liu, Y. Diversity, structure and function of the coiled-coil domains of plant NLR immune receptors. *J. Integr. Plant Biol.* **63**, 283–296 (2021).
19. Baggs, E., Dagdas, G. & Krasileva, K. NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Curr. Opin. Plant Biol.* **38**, 59–67 (2017).
20. Jubic, L. M., Saile, S., Furzer, O. J., El Kasmi, F. & Dangl, J. L. Help wanted: helper NLRs and plant immune responses. *Curr. Opin. Plant Biol.* **50**, 82–94 (2019).
21. Shao, Z. Q. *et al.* Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiol.* **170**, 2095–2109 (2016).
22. Jacob, F., Vernaldi, S. & Maekawa, T. Evolution and conservation of plant NLR functions. *Front. Immunol.* **4**, 1–16 (2013).
23. Saile, S. C. *et al.* Two unequally redundant 'helper' immune receptor families mediate Arabidopsis thaliana intracellular 'sensor' immune receptor functions. *PLOS Biol.* **18**, (2020).
24. Jones, J. D. G., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, 1117 (2016).
25. Baggs, E., Dagdas, G. & Krasileva, K. V. K. V. NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Curr. Opin. Plant Biol.* **38**, 59–67 (2017).
26. Marchal, C., Michalopoulou, V. A., Zou, Z., Cevik, V. & Sarris, P. F. Show me your ID: NLR immune receptors with integrated domains in plants. *Essays Biochem.* **66**, 527–539 (2022).
27. Flor, H. H. Current Status of the Gene-For-Gene Concept. *Annu. Rev. Phytopathol.* **9**, 275–296 (1971).
28. Laflamme, B. *et al.* The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. *Science* **367**, 763–768 (2020).
29. Schreiber, K. J., Chau-Ly, I. J. & Lewis, J. D. What the wild things do: Mechanisms of plant host manipulation by bacterial type III-secreted effector proteins. *Microorganisms* **9**, 1–48 (2021).

30. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the Pseudomonas syringae type III effector HopZ1a. *Proc. Natl. Acad. Sci. USA* **110**, 18722–18727 (2013).
31. Shao, F. *et al.* Cleavage of Arabidopsis PBS1 by a bacterial type III effector. *Science* **301**, 1230–1233 (2003).
32. Mucyn, T. S., Wu, A., Balmuth, A. L., Arasteh, J. M. & Rathjen, J. P. Regulation of Tomato Prf by Pto-like Protein Kinases. **22**, 391–401 (2009).
33. Ntoukakis, V. *et al.* The Tomato Prf Complex Is a Molecular Trap for Bacterial Effectors Based on Pto Transphosphorylation. **9**, (2013).
34. Martin, R. *et al.* Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science* **370**, (2020).
35. Ma, S. *et al.* Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. *Science* **3069**, (2020).
36. Huang, S. *et al.* Identification and receptor mechanism of TIR-catalyzed small molecules in plant immunity. *Science* **377**, (2022).
37. Jia, A. *et al.* TIR-catalyzed ADP-ribosylation reactions produce signaling molecules for plant immunity. *Science* **377**, (2022).
38. Yu, D. *et al.* TIR domains of plant immune receptors are 2′,3′-cAMP/cGMP synthetases mediating cell death. *Cell* **185**, 2370-2386.e18 (2022).
39. Manik, M. K. *et al.* Cyclic ADP ribose isomers: Production, chemical structures, and immune signaling. *Science* **377**, (2022).
40. Wang, J. *et al.* Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science* **364**, eaav5870 (2019).
41. Wang, J. *et al.* Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science* **364**, eaav5868 (2019).
42. Förderer, A. *et al.* A wheat resistosome defines common principles of immune receptor channels. *Nature* **610**, 532–539 (2022).
43. Bi, G. *et al.* The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. *Cell* 1–14 (2021) doi:10.1016/j.cell.2021.05.003.
44. Jacob, P. *et al.* Plant "helper" immune receptors are Ca 2+ -permeable nonselective cation channels. *Science* **373**, 420–425 (2021).
45. Wang, J., Song, W. & Chai, J. Structure, biochemical function, and signaling mechanism of plant NLRs. *Mol. Plant* **16**, 75–95 (2023).
46. Van de Weyer, A. L. *et al.* A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis thaliana. *Cell* **178**, 1260-1272.e14 (2019).
47. Pitblado, R. E. & MacNeill, B. H. Genetic basis of resistance to Pseudomonas syringae pv. tomato in field tomatoes. *Can. J. Plant Pathol.* **5**, 251–255 (1983).
48. Gimenez-Ibanez, S. *et al.* AvrPtoB Targets the LysM Receptor Kinase CERK1 to Promote Bacterial Virulence on Plants. *Curr. Biol.* **19**, 423–429 (2009).
49. Göhre, V. *et al.* Plant Pattern-Recognition Receptor FLS2 Is Directed for Degradation by the Bacterial Ubiquitin Ligase AvrPtoB. *Curr. Biol.* **18**, 1824–1832 (2008).
50. Xiang, T. *et al.* Pseudomonas syringae Effector AvrPto Blocks Innate Immunity by Targeting Receptor Kinases. *Curr. Biol.* **18**, 74–80 (2008).
51. Cheng, W. *et al.* Structural analysis of pseudomonas syringae AvrPtoB bound to host BAK1 reveals two similar kinase-interacting domains in a type III effector. *Cell Host Microbe* **10**, 616–626 (2011).

52. Mathieu, J., Schwizer, S. & Martin, G. B. Pto kinase binds two domains of AvrPtoB and its proximity to the effector E3 ligase determines if it evades degradation and activates plant immunity. *PLoS Pathog.* **10**, 24–27 (2014).
53. Gutierrez, J. R. *et al.* Prf immune complexes of tomato are oligomeric and contain multiple Pto-like kinases that diversify effector recognition. *Plant J.* **61**, 507–518 (2010).
54. Rose, L. E., Michelmore, R. W. & Langley, C. H. Natural variation in the Pto disease resistance gene within species of wild tomato (Lycopersicon). II. Population genetics of Pto. *Genetics* **175**, 1307–1319 (2007).
55. Day, B. *et al.* Molecular basis for the RIN4 negative regulation of RPS2 disease resistance. *Plant Cell* **17**, 1292–1305 (2005).
56. Ade, J., DeYoung, B. J., Golstein, C. & Innes, R. W. Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc. Natl. Acad. Sci. USA* **104**, 2531–2536 (2007).
57. Hurni, S. *et al.* The powdery mildew resistance gene Pm8 derived from rye is suppressed by its wheat ortholog Pm3. *Plant J.* **79**, 904–913 (2014).
58. Jeuken, M. J. W. *et al.* Rin4 causes hybrid necrosis and race-specific resistance in an interspecific lettuce hybrid. *Plant Cell* **21**, 3368–3378 (2009).
59. Chen, C. *et al.* A two-locus interaction causes interspecific hybrid weakness in rice. *Nat. Commun.* **5**, 3357 (2014).
60. Kim, H., Prokchorchik, M. & Sohn, K. H. Investigation of natural RIN4 variants reveals a motif crucial for function and provides an opportunity to broaden NLR regulation specificity. *Plant J.* **110**, 58–70 (2022).
61. Gong, Z. *et al.* The origin and evolution of a plant resistosome. *Plant Cell* **34**, 1600–1620 (2022).
62. Mindrinos, M., Katagiri, F., Yu, G.-L. & Ausubel, F. M. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78**, 1089–1099 (1994).
63. Axtell, M. J. & Staskawicz, B. J. Initiation of RPS2-specified disease resistance in Arabidopsis Is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**, 369–377 (2003).
64. Lewis, J. D., Wu, R., Guttman, D. S. & Desveaux, D. Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. *PLoS Genet.* **6**, e1000894 (2010).
65. Adachi, H. *et al.* Jurassic NLR: Conserved and dynamic evolutionary features of the atypically ancient immune receptor ZAR1. *Plant Cell* 1–50 (2023) doi:10.1093/plcell/koad175.
66. Harant, A., Pai, H., Sakai, T., Kamoun, S. & Adachi, H. A vector system for fast-forward studies of the HopZ-Activated Resistance1 (ZAR1) resistosome in the model plant Nicotiana benthamiana. *Plant Physiol.* **188**, 70–80 (2022).
67. Baudin, M., Hassan, J. A., Schreiber, K. J. & Lewis, J. D. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol.* **174**, 2038–2053 (2017).
68. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the *Pseudomonas syringae*type III effector HopZ1a. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18722–18727 (2013).
69. Harant, A., Pai, H., Sakai, T., Kamoun, S. & Adachi, H. A vector system for fast-forward

studies of the HOPZ-ACTIVATED RESISTANCE1 (ZAR1) resistosome in the model plant Nicotiana benthamiana. *Plant Physiol.* **188**, 70–80 (2022).
70. Schultink, A., Qi, T., Bally, J. & Staskawicz, B. Using forward genetics in Nicotiana benthamiana to uncover the immune signaling pathway mediating recognition of the Xanthomonas perforans effector XopJ4. *New Phytol.* **221**, 1001–1009 (2019).
71. Diplock, N. *et al.* Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1. *Plant. Cell Environ.* **46**, 2238–2254 (2023).
72. Breit-McNally, C., Laflamme, B., Singh, R. A., Desveaux, D. & Guttman, D. S. ZAR1: Guardian of plant kinases. *Front. Plant Sci.* **13**, 1–8 (2022).
73. Mackey, D., Holt, B. F., Wiig, A. & Dangl, J. L. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743–754 (2002).
74. Selote, D. & Kachroo, A. RPG1-B-derived resistance to AvrB-expressing Pseudomonas syringae requires RIN4-like proteins in soybean. *Plant Physiol.* **153**, 1199–1211 (2010).
75. Selote, D. & Kachroo, A. RIN4-like proteins mediate resistance protein-derived soybean defense against Pseudomonas syringae. *Plant Signal. Behav.* **5**, 1453–1456 (2010).
76. Prokchorchik, M. *et al.* A host target of a bacterial cysteine protease virulence effector plays a key role in convergent evolution of plant innate immune system receptors. *New Phytol.* **225**, 1327–1342 (2020).
77. Caldwell, K. S. & Michelmore, R. W. Arabidopsis thaliana genes encoding defense signaling and recognition proteins exhibit contrasting evolutionary dynamics. *Genetics* **181**, 671–684 (2009).
78. Albers, P., Ustun, S., Witzel, K., Kraner, M. & Bornke, F. A remorin from Nicotiana benthamiana interacts with the Pseudomonas type-III effector protein HopZ1a and is phosphorylated by the immune-related kinase PBS1. *BioRxiv* 1–49 (2018).
79. Bai, R. *et al.* Decoy engineering of the receptor-like cytoplasmic kinase StPBS1 to defend against virus infection in potato. *Mol. Plant Pathol.* **23**, 901–908 (2022).
80. Carter, M. E. *et al.* Convergent evolution of effector protease recognition by arabidopsis and barley. *Mol. Plant-Microbe Interact.* **32**, 550–565 (2019).
81. Sun, J. *et al.* Comparative study of Arabidopsis PBS1 and a wheat PBS1 homolog helps understand the mechanism of PBS1 functioning in innate immunity. *Sci. Rep.* **7**, 1–12 (2017).
82. Pilet-Nayel, M. L. *et al.* Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Front. Plant Sci.* **8**, 1–9 (2017).
83. Dong, O. X. *et al.* Marker-free carotenoid-enriched rice generated through targeted gene insertion using CRISPR-Cas9. *Nat. Commun.* **11**, 1–10 (2020).
84. Wilson, I. W., Schiff, C. L., Hughes, D. E. & Somerville, S. C. Quantitative Trait Loci Analysis of Powdery Mildew Disease Resistance in the Arabidopsis thaliana Accession Kashmir-1. *Genetics* **158**, 1301–1309 (2001).
85. Weigel, D. & Mott, R. The 1001 Genomes Project for *Arabidopsis thaliana*. *Genome Biol.* **10**, 107 (2009).
86. Aranzana, M. J. *et al.* Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genet.* **1**, e60 (2005).
87. Aoun, N. *et al.* Quantitative Disease Resistance under Elevated Temperature: Genetic Basis of New Resistance Mechanisms to Ralstonia solanacearum. *Front. Plant Sci.* **8**, 1–

16 (2017).
88. Huard-Chauveau, C. *et al.* An Atypical Kinase under Balancing Selection Confers Broad-Spectrum Disease Resistance in Arabidopsis. *PLoS Genet.* **9**, e1003766 (2013).
89. Wang, G. *et al.* The decoy substrate of a pathogen effector and a pseudokinase specify pathogen-induced modified-self recognition and immunity in plants. *Cell Host Microbe* **18**, 285–295 (2015).
90. Baudin, M. *et al.* A natural diversity screen in Arabidopsis thaliana reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex. *Plant. Cell Environ.* **44**, 629–644 (2021).
91. Barragan, C. A. *et al.* RPW8/HR repeats control NLR activation in Arabidopsis thaliana. *PLoS Genet.* **15**, 1–21 (2019).
92. Chen, C., Zhiguo, E. & Lin, H. X. Evolution and molecular control of hybrid incompatibility in plants. *Front. Plant Sci.* **7**, 1–10 (2016).
93. Calvo-Baltanás, V., Wang, J. & Chae, E. Hybrid Incompatibility of the Plant Immune System: An Opposite Force to Heterosis Equilibrating Hybrid Performances. *Front. Plant Sci.* **11**, (2021).
94. Montes, N., Alonso-Blanco, C. & García-Arenal, F. *Cucumber mosaic virus* infection as a potential selective pressure on *Arabidopsis thaliana* populations. *PLOS Pathog.* **15**, e1007810 (2019).
95. Glander, S. *et al.* Assortment of flowering time and immunity alleles in natural Arabidopsis thaliana populations suggests immunity and vegetative lifespan strategies coevolve. *Genome Biol. Evol.* **10**, 2278–2291 (2018).
96. MacQueen, A., Sun, X. & Bergelson, J. Genetic architecture and pleiotropy shape costs of Rps2-mediated resistance in Arabidopsis thaliana. *Nat. Plants* **2**, 16110 (2016).
97. Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M. & Bergelson, J. Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana. *Nature* **423**, 74–77 (2003).
98. Karasov, T. L. *et al.* The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* **512**, 436–440 (2014).
99. Lehti-Shiu, M. D. & Shiu, S.-H. S. H. Diversity, classification and function of the plant protein kinase superfamily. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2619–2639 (2012).
100. Ellis, J. G., Lawrence, G. J., Luck, J. E. & Dodds, P. N. Identification of Regions in Alleles of the Flax Rust Resistance Gene L That Determine Differences in Gene-for-Gene Specificity. *Plant Cell* **11**, 495 (1999).
101. 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.* **103**, 8888–8893 (2006).
102. Martel, A. *et al.* Immunodiversity of the Arabidopsis ZAR1 NLR Is conveyed by receptor-like cytoplasmic kinase sensors. *Front. Plant Sci.* **11**, 1290 (2020).
103. Liu, C. *et al.* Two Arabidopsis receptor-like cytoplasmic kinases SZE1 and SZE2 associate with the ZAR1–ZED1 complex and are required for effector-triggered immunity. *Mol. Plant* **12**, 967–983 (2019).
104. Bastedo, D. P. *et al.* Perturbations of the ZED1 pseudokinase activate plant immunity. *PLoS Pathog.* **15**, e1007900 (2019).
105. Stephens, C., Hammond-Kosack, K. E. & Kanyuka, K. WAKsing plant immunity, waning diseases. *J. Exp. Bot.* **73**, 22–37 (2022).
106. Feng, F. *et al.* A Xanthomonas uridine 5′-monophosphate transferase inhibits plant

immune kinases. *Nature* **485**, 114–118 (2012).
107. Qi, D. *et al.* Recognition of the protein kinase AVRPPHB SUSCEPTIBLE1 by the disease resistance protein RESISTANCE TO PSEUDOMONAS SYRINGAE5 Is dependent on S-acylation and an exposed loop in AVRPPHB SUSCEPTIBLE. *Plant Physiol.* **164**, 340–351 (2014).
108. Kim, S. H., Qi, D., Ashfield, T., Helm, M. & Innes, R. W. Using decoys to expand the recognition specificity of a plant disease resistance protein. *Science* **351**, 684–687 (2016).
109. Pottinger, S. E. *et al.* Optimizing the PBS1 decoy system to confer resistance to potyvirus infection in Arabidopsis and soybean. *Mol. Plant-Microbe Interact.* **33**, 932–944 (2020).
110. Helm, M. *et al.* Engineering a decoy substrate in soybean to enable recognition of the soybean mosaic virus NIa protease. *Mol. Plant-Microbe Interact.* **32**, 760–769 (2019).
111. Chen, J. *et al.* Loss of AvrSr50 by somatic exchange in stem rust leads to virulence for Sr50 resistance in wheat. *Science* **358**, 1607–1610 (2017).
112. Tamborski, J., Seong, K., Liu, F., Staskawicz, B. J. & Krasileva, K. V. Altering Specificity and Autoactivity of Plant Immune Receptors Sr33 and Sr50 Via a Rational Engineering Approach. *Mol. Plant-Microbe Interact.* **36**, 434–446 (2023).
113. Salcedo, A. *et al.* Variation in the AvrSr35 gene determines Sr35 resistance against wheat stem rust race Ug99. *Science* **358**, 1604–1606 (2017).
114. Cesari, S. *et al.* The rice resistance protein pair RGA4/RGA5 recognizes the Magnaporthe oryzae effectors AVR-Pia and AVR1-CO39 by direct binding. *Plant Cell* **25**, 1463–1481 (2013).
115. Ashikawa, I. *et al.* Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer Pikm-specific rice blast resistance. *Genetics* **180**, 2267–2276 (2008).
116. Okuyama, Y. *et al.* A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. *Plant J.* **66**, 467–479 (2011).
117. Maqbool, A. *et al.* Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. *Elife* **4**, 1–24 (2015).
118. Ortiz, D. *et al.* Recognition of the Magnaporthe oryzae Effector AVR-Pia by the Decoy Domain of the Rice NLR Immune Receptor RGA5. *Plant Cell* **29**, 156–168 (2017).
119. Guo, L. *et al.* Specific recognition of two MAX effectors by integrated HMA domains in plant immune receptors involves distinct binding surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 11637–11642 (2018).
120. Liu, Y. *et al.* A designer rice NLR immune receptor confers resistance to the rice blast fungus carrying noncorresponding avirulence effectors. *Proc. Natl. Acad. Sci. U. S. A.* **118**, 1–9 (2021).
121. Césari, S. *et al.* New recognition specificity in a plant immune receptor by molecular engineering of its integrated domain. *Nat. Commun.* **13**, 1–13 (2022).
122. Zhang, X. *et al.* The synthetic NLR RGA5 HMA5 requires multiple interfaces within and outside the integrated domain for effector recognition. 1–13 (2024) doi:10.1038/s41467-024-45380-2.
123. De la Concepcion, J. C. *et al.* Protein engineering expands the effector recognition profile of a rice NLR immune receptor. *Elife* **8**, 611152 (2019).
124. Maidment, J. H. R. *et al.* Effector target-guided engineering of an integrated domain expands the disease resistance profile of a rice NLR immune receptor. *Elife* **12**, 1–25

(2023).
125. Bao, G., Tang, M., Zhao, J. & Zhu, X. Nanobody: a promising toolkit for molecular imaging and disease therapy. *EJNMMI Res.* **11**, (2021).
126. Kourelis, J., Marchal, C., Posbeyikian, A., Harant, A. & Kamoun, S. NLR immune receptor–nanobody fusions confer plant disease resistance. *Science* **379**, 934–939 (2023).

# CHAPTER II: A natural diversity screen in *Arabidopsis thaliana* reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex

## PREFACE

This chapter was published in 2021 as "A natural diversity screen in *Arabidopsis thaliana* reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex" in *Plant, Cell, & Environment*. I performed the western blot assays and provided feedback on manuscript. The publication has been reproduced below with the permission of all authors.

This work uses natural diversity to identify determinants for ZAR1 recognition of the *P. syringae* effector HopZ1a, which is a key step towards rational design of ZAR1 recognition pathways for *P. syringae* resistance.

Baudin M., Martin E.C., Sass C., Hassan J.A., Bendix C., Sauceda R., Diplock N., Specht C.D., Petrescu A.J., Lewis J.D. A natural diversity screen in *Arabidopsis thaliana* reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex. *Plant Cell Environ* (44) 629-644 (2021).

**Abstract**

Pathogen pressure on hosts can lead to the evolution of genes regulating the innate immune response. By characterizing naturally occurring polymorphisms in immune receptors, we can better understand the molecular determinants of pathogen recognition. ZAR1 is an ancient *Arabidopsis thaliana* NLR (Nucleotide-binding [NB] Leucine-rich-repeat [LRR] Receptor) that recognizes multiple secreted effector proteins from the pathogenic bacteria *Pseudomonas syringae* and *Xanthomonas campestris* through its interaction with receptor-like cytoplasmic kinases (RLCKs). ZAR1 was first identified for its role in recognizing *P. syringae* effector HopZ1a, through its interaction with the RLCK ZED1. To identify additional determinants of HopZ1a recognition, we performed a computational screen for ecotypes from the 1001 genomes project that were likely to lack HopZ1a recognition, and tested ~300 ecotypes. We identified ecotypes containing polymorphisms in ZAR1 and ZED1. Using our previously established *Nicotiana benthamiana* transient assay and Arabidopsis ecotypes, we tested for the effect of naturally occurring polymorphisms on ZAR1 interactions and the immune response. We identified key residues in the NB or LRR domain of ZAR1 that impact the interaction with ZED1. We demonstrate that natural diversity combined with functional assays can help define the molecular determinants and interactions necessary to regulate immune induction in response to pathogens.

**Introduction**

Variation among individuals in natural populations is a powerful tool to examine the importance of natural selection on phenotypes of interest, and such studies of phenotypic diversity have been expedited by advances in sequencing technologies and resequencing of ecotypes of a species. Natural genetic diversity occurring in Arabidopsis has been successfully used to characterize major determinants of autoimmunity and of immunity to pathogenic bacteria or viruses [1–6]. The 1001 genomes project generated whole genome sequencing data for 1,135 Arabidopsis accessions, which has enabled the detection of naturally occurring polymorphisms associated with phenotypes of interest [7–10]. Since pathogen pressure on hosts can lead to the

evolution of genes regulating the innate immune response, the characterization of naturally occurring polymorphisms can help us to understand the molecular mechanisms associated with plant pathogen recognition.

Plant genomes encode large families of protein receptors that can recognize pathogen features and activate disease resistance [11–15]. Intracellular Nucleotide-binding [NB] Leucine-rich repeat [LRR] Receptors (NLRs) specifically recognize pathogen-secreted proteins called effector proteins [16–18]. Effector-triggered immunity (ETI) mediated by NLRs induces a specific and rapid immune response often resulting in a form of localized cell-death called the hypersensitive response (HR) [17,19,20]. NLRs can detect effectors by direct interaction as proposed by the receptor-ligand model [21,22]. However, for most effector proteins from the bacterial pathogen *Pseudomonas syringae*, NLRs indirectly recognize effector proteins through their activity on a virulence target of the effector (guard model) or a decoy of this target (decoy model) [16,17,23]. Modification of the guarded protein then triggers NLR activation which induces defense responses. Some NLRs contain an extra domain called a sensor domain or integrated decoy that can directly sense the effector [16,24–26].

NLRs generally contain three main protein domains: a variable N-terminal domain that is often a CC (Coiled-Coil) or TIR (Toll/Interleukin1 Receptor) domain, followed by a central NBARC domain (Nucleotide-Binding Apaf1-R protein CED4) and C-terminal Leucine-Rich Repeat (LRR) domain. The central NBARC domain can be further subdivided into 3 subdomains (NBD [nucleotide-binding domain], HD1 [helical domain 1] and the WHD [winged-helix domain]). Detailed mechanistic studies and the resolution of individual domains as well as full-length NLR structures led to a molecular switch model for the activation of NLRs [27–30]. In the absence of the cognate effector, NLRs are kept in an "off" ADP-bound state by intramolecular interactions. Effector detection often occurs (directly or indirectly) at the C-terminal LRR domain, and results in the exchange of ADP for ATP in the NBARC domain due to deep conformational reorganization of NLR domains. In many cases, oligomerization of the N-terminal domain (CC or TIR) is associated with the activation of the NLR and the induction of downstream signaling [31–34]. The mechanisms by which activated NLRs induce defense responses remain elusive as very few signaling pathways have been identified so far.

The NLR HOPZ-ACTIVATED RESISTANCE 1 (ZAR1) recognizes multiple unrelated bacterial effector proteins from two different bacterial pathogens [35–40], through a strategy that is consistent with the decoy model [41]. ZAR1 was first characterized in Arabidopsis as being required for the induction of the HR to HopZ1a, a type III secreted effector (T3SE) from the pathogenic bacteria *P. syringae*, as well as a restriction in bacterial growth of strains carrying HopZ1a [37,38,42]. In addition, ZAR1 is required for the recognition of AvrAC, a T3SE from *Xanthomonas campestris*, XopJ4, a member of the HopZ/YopJ family of T3SEs from *Xanthomonas performans*, and HopF2a, HopO1c, HopX1i and HopBA1a, T3SEs from *P. syringae* [35,36,39,40,43]. We previously demonstrated that Arabidopsis ZAR1 can complement the down-regulation of *ZAR1* from *Nicotiana benthamiana* to recognize HopZ1a, suggesting that ZAR1 is functionally conserved between the Brassicaceae and the Solanaceae [34].

ZAR1 interacts with the pseudokinase HOPZ-EFFECTOR-TRIGGERED IMMUNITY DEFICIENT (ZED1) which is acetylated by HopZ1a [34,44,45]. The modification of ZED1 by HopZ1a is hypothesized to trigger ZAR1 activation [44]. ZED1 is a receptor-like cytoplasmic kinase (RLCK) that is located in a genomic cluster of kinase-encoding genes (called ZRKs for ZED1-RELATED KINASES), some of which have been shown to interact with ZAR1 [39,44]. ZAR1-dependent AvrAC and HopF2a recognition is mediated by ZRK1 (also called Resistance

related KinaSe 1, RKS1) and ZRK3 respectively (Seto et al., 2017; G. Wang et al., 2015). HopO1c, HopX1i and HopBA1a require ZRK3, ZED1 and ZRK2 respectively for recognition [46]. In *Nicotiana benthamiana*, ZAR1 interacts with a kinase, XOPJ4 IMMUNITY 2 (JIM2) to recognize XopJ4 [36,43].

We previously showed that protein domains of ZAR1 are involved in specific intra- and intermolecular interactions that are critical for ZAR1 activity and prevent inappropriate activation of the NLR [33,34,47]. The LRR domain of ZAR1 is important for interactions with ZED1, RKS1/ZRK1 and ZRK3 (Baudin et al., 2017; G. Wang et al., 2015). Oligomerization of $ZAR1^{CC}$ is associated with the induction of HR [34] and is required for ZAR1 activity [33,48,49]. Based on recent cryo-EM structures of ZAR1, it has been hypothesized that upon activation, ZAR1 oligomers form a pore in the plasma membrane and induce cell death [48,49]. The cryo-EM structures of ZAR1 in complex with two kinases, RKS1 and PBL2, in the inactive and active states [29,48] provide excellent tools to study the conformational dynamics of this NLR.

Here we mined data from the 1001 genomes project to investigate genetic diversity in HopZ1a recognition across Arabidopsis ecotypes. We identified polymorphisms primarily in ZAR1, and used complementation and interaction assays to further investigate the impact of these polymorphisms on ZAR1 function. We analyzed the polymorphisms together with previously identified EMS mutations in the context of ZAR1 cryo-EM structures. This work combines the power of natural diversity with structural analysis to define the molecular determinants and interactions necessary to regulate the induction of immunity in response to HopZ1a.

## Materials and Methods

*Plasmid construction*

Phusion polymerase (New England Biolabs, Ipswich, MA, USA) was used for all cloning, and constructs were confirmed by sequencing. Sequence analysis was performed with CLC Main Workbench. *ZAR1* alleles were cloned from genomic DNA isolated from their respective ecotypes and single mutations were generated using the Q5 Site-directed mutagenesis kit (New England Biolabs). *ZAR1* variants were amplified by PCR to add a 5'*Xho*I site and cloned into the pMac15 vector to maintain the frame for the vector-encoded C-terminal 5xMyc tag. The pMac15 vector was modified from pBD to contain a 5xMyc tag between the *Stu*I and *Spe*I sites, as described in [34]. The *ZED1* 3xFlag construct was cloned by a crossover PCR approach to add the 3xFlag tag in-frame at the 3′end. The PCR products were cloned into the pMac14 vector, modified from pBD [37,50]. The *hopZ1a* constructs were described in [37]. The *ZAR1* constructs used for the genetic transformation of Arabidopsis was designed by amplifying a genomic fragment of *ZAR1* containing a ~1kb of the promoter region and the coding sequence, or the coding sequence and a 5xMyc tag. For the promoter-ZAR1-5xMyc construct, the fragments were amplified by PCR to add a 5' *Sal*I site and an attB1 site, and a 3' *Not*I site and an attB2 site. The resulting promoter-ZAR1-myc product was first cloned into pBS in a three-body ligation, followed by a BP reaction into pDONR207, and a LR reaction into linearized pEarleyGate303 [51]. The VIGS constructs were described in [34]. In brief, we cloned a genomic fragment of *NbZAR1* gene, targeting ~700 bp region at the 5'end of the gene, into the pYL156 vector [52]. Oligonucleotides used for cloning are listed in Table S1.

*Plant material and plant transformation*

*Arabidopsis thaliana* seeds were vapor-sterilized [53] and germinated on 0.5X Murashige and Skoog and 0.8% agar media. After germination, the seedlings were transplanted on Sunshine #1 soil (lacking RSi) supplemented with 20:20:20 fertilizer and grown in a growth chamber at 22°C with 9-h-light (~130 $\mu Em^{-2}$ $s^{-1}$) and 15-h-dark cycles. To induce flowering, IP-deh was vernalized for ~12 weeks in the cold room (4°C, 8-h-light ~130 $\mu Em^{-2}$ $s^{-1}$ and 16-h-dark cycles) and then transferred to long day chambers (22°C, 16-h-light ~130 $\mu Em^{-2}$ $s^{-1}$ and 8-h-dark cycles). Cvi-0 and *zar1-1* were moved to long day chambers (without vernalization) to induce flowering. Arabidopsis transgenic plants were generated using the floral dip method with *Agrobacterium tumefaciens* strain GV3101 [53]. Transgenic plants were selected by Basta resistance in Sunshine #1 soil supplemented with 20:20:20 fertilizer. Transgenes were confirmed by PCR and the expression of ZAR1-5xMyc was tested by western blot analysis. The genetics of T3 transgenic plants were determined by segregation ratios for Basta resistance. For *Nicotiana benthamiana*, plants were prepared and transiently transformed as previously described in [33]. Briefly, *A. tumefaciens* GV2260 cultures were grown overnight at 28°C in LB broth with kanamycin and rifampicin. The next day, the bacteria were resuspended in 1 ml of 10 mM MES, pH 5.6 and 200 µM of acetosyringone and incubated in the dark for ~4h. The cultures were then adjusted to an optical density (λ = 600 nm, OD600) of 0.25 per construct. The underside of the leaves of 5- to 7-week-old *N. benthamiana* plants were infiltrated by hand with a needleless syringe. For dexamethasone-inducible constructs, the plants were sprayed with 20 µM dexamethasone (Sigma-Aldrich) 12–24 h after inoculation. Tissue collection or observations were performed from 3 to 48 h after dexamethasone induction depending on the experiment. For the VIGS experiments, the plants were used 1–2 weeks after inoculation.

*Plant infection assays and conductivity*

HR assays in Arabidopsis were performed by infiltrating *Pto*DC3000 strains carrying *hopZ1a* or empty pUCP20 vector (*ev*). Bacteria were resuspended to an OD600 = 0.1 (~$5x10^7$ cfu/mL) in 10 mM $MgCl_2$ for HR assays. Diluted inocula were hand infiltrated by using a needleless syringe as described by [54]. The HR was scored at 16 to 24 hpi. For ion leakage assays in *N. benthamiana*, two disks (1.5 $cm^2$) were harvested, soaked in nanopure water for 1 h, and transferred to 6 mL of nanopure water. Readings were taken with an Orion 3 Star conductivity meter (Thermo Electron). The conductivity data were analyzed using Minitab 19 software (Minitab) by a one-way ANOVA statistical test at a significance level of P = 0.05 followed by multiple comparisons of means using Tukey's post hoc test.

*Sequence alignment, tree-building, and sequence logos*

ZAR1 and ZED1 protein and nucleotide sequences for Arabidopsis ecotypes were obtained from the 1001 Genomes Project website (http://signal.salk.edu/atg1001/index.php). Sequences from 855 ecotypes were obtained for this analysis. The 1001 Genomes Project website hosted by the Max Planck institute (http://1001genomes.org/) contains the latest releases, but did not provide protein data at the time. Protein and nucleotide sequences were aligned using MAFFT version 7 [55], and protein sequences were used to verify and align nucleotide sequences. Once aligned, nucleotide sequences were grouped by 100% identity: Sequences containing single base pair changes which did not change the encoded amino acid and were otherwise identical to the rest of a group were included in a group. Sequences with premature stop codons were placed in their own unique groups. For tree-building, one representative sequence from each group with identical sequences was selected. These filtered sequences were re-aligned and subjected to

phylogenetic analysis. Trees were generated with MrBayes [56] partitioned by codon using GTR+G using 100 million generations, sampled every 1000 generations. Branches with less than 50 percent posterior probability were collapsed.

*Identification of candidate ecotypes*

HR assays were performed on 91 ecotypes of the Nordborg collection [57], and classified as ETI-sufficient or ETI-deficient. Based on the sequence groups for the two phenotypic categories, additional ecotypes with ZAR1 sequences similar to those in the ETI-deficient group were identified from the 1001 Genomes collection. Strong candidate ETI-deficient ecotypes were those that were unique (not part of a sequence identity group), in a group with other ETI-deficient ecotypes, or in a group with few ETI-sufficient ecotypes. Sequences in a large group with many ETI-sufficient ecotypes were considered as weak candidates. Next, the ZED1 sequence identity groups were added into the analysis, with unique sequences and small groups preferred. Finally, ETI-sufficient groups for both genes (*ZAR1* and *ZED1*) were considered. The strongest candidates in the *ZAR1*/*ZED1* category were ETI-deficient ecotypes that had no overlap with ETI-sufficient ecotypes, especially if their groups for one of the genes were also unique or small. ETI-deficient ecotypes with identical *ZAR1*/*ZED1* sequences as ETI-sufficient ecotypes were removed from the analysis. Next, the sequence identity groups were analyzed together with the *ZAR1* and *ZED1* gene trees. Groups containing ETI-deficient ecotypes were identified on the tree, and the *ZAR1* and *ZED1* sequences were extracted for related ecotypes as well as the other ecotypes in the ETI-deficient sequence identity group. Each sequence identity group was annotated according to whether an ETI-sufficient ecotype was present within the group. Ideal candidate ecotypes were either unique, in small groups that did not contain ETI-sufficient ecotypes, or were closely related to an ETI-deficient ecotype. Ecotypes that were present only in ETI-sufficient groups for both genes were removed. We selected all ecotypes that fulfilled these criteria and contained a unique sequence for *ZAR1* and/or *ZED1*, as well as representatives from smaller sequence identity groups. This approach resulted in a list of 202 ecotypes that were tested in the secondary screen. Since a few of the 1001 genomes seed stocks were later found to contain errors [58], we removed affected ecotypes from the analysis.

*Gene expression analyses*

Total RNA was extracted from Arabidopsis leaves using TRIzol reagent (Ambion). First-strand cDNA synthesis was performed using 1 µg of total RNA with an anchored oligo(dT) using SuperScript III reverse transcriptase (Invitrogen). DNA was removed using DNAse I treatment (Invitrogen). Quantitative RT-PCR was performed on 96-well plates, with the BIO-RAD C1000 system and using the iQ™ SYBR® Green Supermix (BIO-RAD). Transcript levels for each of the target genes were normalized to the endogenous *TIP41* transcript level [59] using $2^{-\Delta\Delta CT}$ method [60]. The data shown are mean values obtained from three independent biological experiments with three technical repeats.

*Co-immunoprecipitation*

In planta co-immunoprecipitation was performed as described in [33] with some slight modifications. A total of 3 $cm^2$ of *N. benthamiana* leaves transiently expressing our genes of interest was used. Tissue was homogenized in liquid nitrogen, and protein complexes were extracted using 1 ml of IP1 buffer (50 mM HEPES, 50 mM NaCl, 10 mM ethylene diamine tetraacetic acid (EDTA), 0.2% TritonX-100, and 0.1 mg $ml^{-1}$ Dextran [Sigma-Aldrich D1037, St.

Louis, MO, USA], pH 7.5). The clarified extract was then mixed with 5 µl of Myc-Trap® magnetic beads [Chromotek] for 3 h at 4°C. The immunoprecipitated proteins were washed six times with IP2 buffer (50 mM HEPES, 150 mM NaCl, 10 mM EDTA, and 0.1% Triton X-100, pH 7.5) on a magnetic stand. IP1 and IP2 buffers were supplemented with proteinase inhibitor cocktail (1 mM PMSF, 1 mg $ml^{-1}$ leupeptin [Sigma-Aldrich L2023], 1mg $ml^{-1}$ aprotinin [Sigma-Aldrich A6191], 1 mg $ml^{-1}$ antipain [Sigma-Aldrich A1153], 1 mg $ml^{-1}$ chymostatin [Sigma-Aldrich C7268], and 1 mg $ml^{-1}$ pepstatin [Sigma-Aldrich P5315]). The immunoprecipitated fraction was eluted by boiling the bead in Laemmli buffer. Finally, the input and immunoprecipitated fractions were separated on 12% acrylamide gels by SDS- PAGE and detected by immunoblot using anti-c-Myc antibody (Abcam ab32072, Cambridge, MA, USA), or anti-Flag antibody (Sigma-Aldrich F1804) followed by the appropriate secondary antibody coupled with horseradish peroxidase.

**Results**

*New ZAR1 and ZED1 polymorphisms identified in a natural diversity screen*

To investigate how the HopZ1a-induced immune response is conserved among Arabidopsis ecotypes, we tested 91 accessions from the Nordborg collection [57], for their ability to recognize HopZ1a and elicit ETI. We infiltrated three or four individuals from each ecotype with *P. syringae* pv. *tomato* (*Pto*) DC3000 carrying an empty pUCP20 vector (*Pto*DC3000 (*ev*)), as a negative control, or *Pto*DC3000 carrying HopZ1a (*Pto*DC3000 (*hopZ1a*)), and phenotyped the plants for the HR 16-20 hours post-infiltration. *Pto*DC3000 carrying HopZ1a has previously been shown to induce a strong and rapid HR that is dependent on ZAR1 and ZED1 in *A. thaliana* Col-0 and is associated with a restriction in bacterial growth [37,38,44]. We grouped ecotypes into two categories, based on the presence of the HopZ1a-induced HR. We classified ecotypes with no HR or weak delayed HR in response to HopZ1a as ETI-deficient, and ecotypes with strong variable HR or very strong HR as ETI-sufficient. Among the 91 ecotypes tested, 79 were ETI-sufficient and 12 were ETI-deficient, indicating that 13.2% of the Nordborg collection did not induce a strong and rapid HR in response to HopZ1a (Figure 1A).

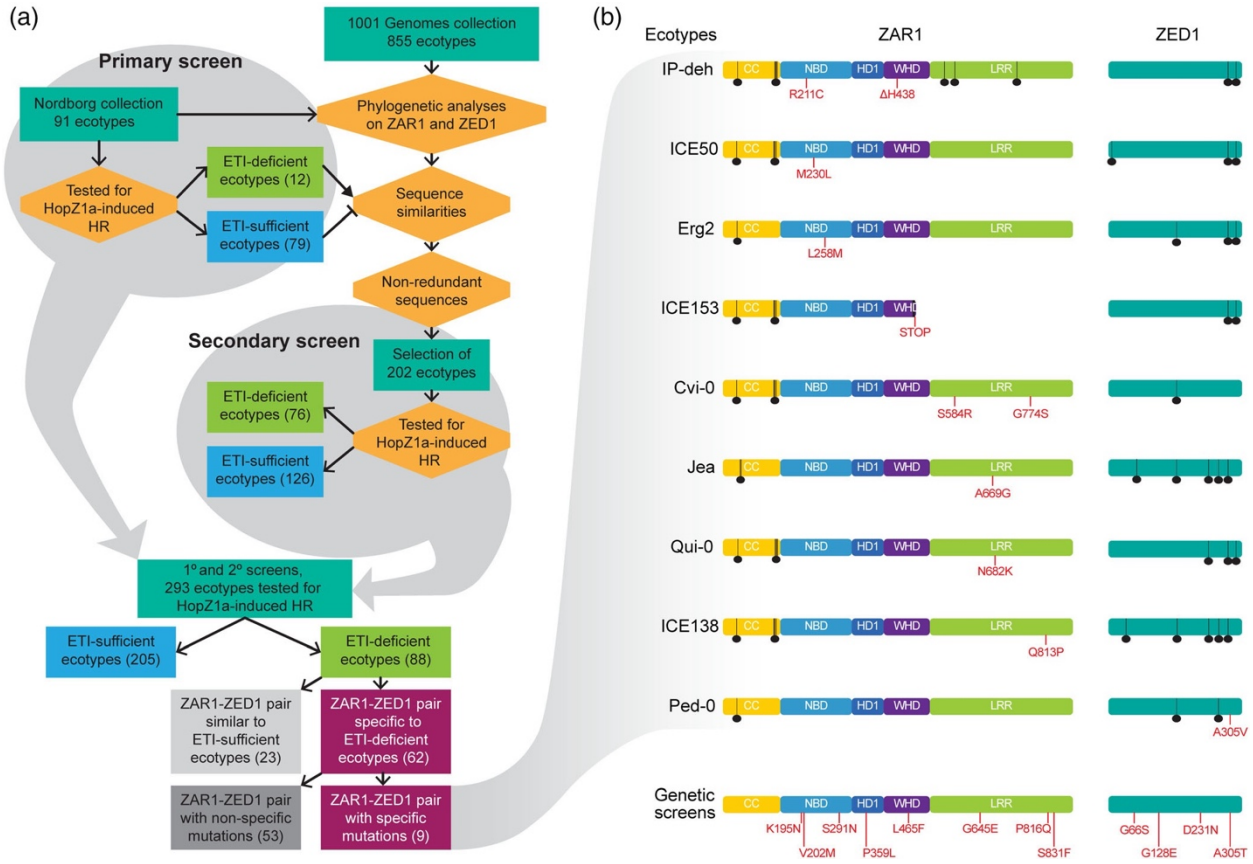



**FIGURE 1: Arabidopsis natural diversity screen.** A. Schematic of the workflow used to identify and test 297 ecotypes from the Nordborg and 1001 genome collections. ETI-deficient ecotypes are not able to recognize HopZ1a, while ETI-sufficient ecotypes recognize HopZ1a. The number of ecotypes identified for each category is indicated in parentheses. B. ZAR1 and ZED1 schematics with polymorphisms identified in the nine ecotypes from the natural diversity screen as well as in the previous genetic screens [39,44]. The colored boxes correspond to the following domains: yellow is the CC domain, medium blue is the nucleotide-binding domain (NBD) subdomain, dark blue is the helical domain 1 (HD1) subdomain, purple is the winged helix domain (WHD) subdomain, and green is the leucine-rich repeat (LRR) domain. The amino acid changes induced by the specific polymorphisms are indicated below the schematic in red. Polymorphisms that differ for Col-0 sequences and are present in ETI-sufficient ecotypes are indicated with black dots.

We took advantage of the resources available at the 1001 genomes project to investigate *ZAR1* and *ZED1* diversity among the 855 ecotype genome sequences available at the time. We retrieved and aligned *ZAR1* and *ZED1* sequences from the 855 ecotypes (http://1001genomes.org/) and performed phylogenetic analyses to produce independent gene trees. Multiple clades were recovered with low bootstrap support, thus clear differentiation and delimitation of groups was not always possible. The ETI-deficient ecotypes were fairly evenly distributed across the ZAR1 and ZED1 gene trees, indicating that the loss of HopZ1a recognition occurred independently (Figure S1). To identify additional ecotypes from the 1001 genomes project that were likely to be ETI-deficient, we grouped ecotypes by the similarity of their *ZAR1* or *ZED1* sequences based on the phylogenetic analysis and their responses in the primary screen on the Nordborg collection. We prioritized ecotypes if the *ZAR1* (or *ZED1*) sequences were a) unique (not similar to any other groups), b) in a group with other ETI-deficient ecotypes, or c) in a group with a few ecotypes showing ETI (Figure 1A). We included the two latter groups in order to maximize our chances of obtaining ETI-deficient *ZAR1* (or *ZED1*) sequences. We then

compared the sequences of *ZAR1* and *ZED1* in the ETI-deficient groups versus the ETI-sufficient groups. We classified ecotypes as strong candidates for ETI-deficiency if the *ZAR1/ZED1* sequences from ecotypes in ETI-deficient groups were different from those in ETI-sufficient groups. These analyses yielded 202 ecotypes for which we tested three or four individuals of each ecotype for the HopZ1a-induced HR. The ecotypes were binned into ETI-deficient or ETI-sufficient categories depending on their capacity to recognize HopZ1a (Figure 1A). Out of the 202 accessions functionally tested in the secondary screen, we identified 76 ecotypes (38%) that were affected in the recognition of HopZ1a, which was substantially enriched compared to our primary screen. Altogether from both screens, we tested nearly 300 Arabidopsis ecotypes and found 88 ecotypes (30%) that lacked the capacity to recognize HopZ1a.

Based on our two screens, we re-generated *ZAR1* and *ZED1* consensus sequences for ecotypes in each HR category and looked for polymorphisms in ZAR1 and ZED1 that could explain phenotypic differences in the HR. Overall, we saw very high conservation of ZAR1 and ZED1. Among the 88 ETI-deficient ecotypes, 23 had an identical ZAR1-ZED1 pair as ecotypes in the ETI-sufficient category. Therefore, the lack of HopZ1a recognition in these 23 ecotypes is likely due to ZAR1- and/or ZED1-independent polymorphisms. In addition, three ecotypes (Fi-0, BEZ-9 and ARR-17) had seed issues and were removed from the analysis [58]. Among the remaining 62 ETI-deficient ecotypes, 53 had a specific ZAR1-ZED1 pair. However either ZAR1 or ZED1 had polymorphisms that were found in the ETI-sufficient ecotypes, and therefore were not unique to an ETI-deficient phenotype. While the specific combination of ZAR1 and ZED1 polymorphisms present in these ecotypes could be responsible for the ETI-deficient phenotype, we focused on polymorphisms that were only found in ETI-deficient ecotypes. That left us with nine ecotypes carrying polymorphisms in ZAR1 or ZED1 that were unique to the ETI-deficient category (Figure 1A). A total of 10 unique ETI-deficient specific polymorphisms (1 STOP codon, 1 amino acid deletion and 8 substitutions) were found in ZAR1 sequences across 8 different ecotypes (Figure 1B). The identification of the STOP codon in the ICE153 ecotype was a good validation of the screen, as we previously showed that truncating the $ZAR1^{LRR}$ domain strongly impaired ZAR1 function [34]. For ZED1, we identified a single substitution that was specific to the ETI-deficient category, A305V found in the Ped-0 accession. Interestingly, the A305 residue in ZED1 was previously identified in a genetic screen as being required for HopZ1a-induced immunity [34,44]. We cloned *ZAR1* and *ZED1* out of the nine ecotypes identified during the screen and found minor differences between the actual cloned sequence and the sequencing information available in the 1001 genomes database, partly due to missing bases in the database (Figure S2, S3). To ensure that short-read sequence quality did not impact our screen, we analyzed recently generated long-read NLR sequence data from Van de Weyer and collaborators on 64 accessions [6]. This data set would not have been sufficient for our screen as only 23 ecotypes overlapped with the ~300 ecotypes we screened, and only two ecotypes were included in the nine ecotypes we identified. In addition, the sequence data from Van de Weyer and collaborators focused on NLRs and did not include genes such as *ZED1* [6]. Based on our screen of ~300 ecotypes, we focused this study on nine ecotypes containing ZAR1-ZED1 pairs with eleven polymorphisms that might account for the loss of HopZ1a-induced immunity (Figure 1B).

*ZAR1 and ZED1 mRNA level in ETI-deficient ecotypes*

To verify that the phenotype observed in the nine ecotypes was not the consequence of reduced gene expression of *ZAR1* and *ZED1*, we measured the mRNA levels of the two genes by

qRT-PCR in each ecotype. The primers were designed to target conserved sequences across all nine ecotypes and to be specific for *ZAR1* or *ZED1* genes using primer blast [61]. The relative expression of *ZAR1* in the ecotypes was statistically the same as Col-0 (Figure 2A). However, we observed higher expression of *ZED1* in all the ecotypes (4- to 5-fold greater) relative to Col-0 (Figure 2B). Interestingly the Ped-0 ecotype, which carries the A305V polymorphism in ZED1, had 9-fold higher expression of *ZED1* compared to *ZED1* expression in Col-0 (Figure 2B). It is also important to note that none of the ecotypes showed reduced expression of *ZAR1* or *ZED1*, which excludes the possibility that low levels of *ZAR1* and/or *ZED1* are the cause of the ETI-deficient phenotype.

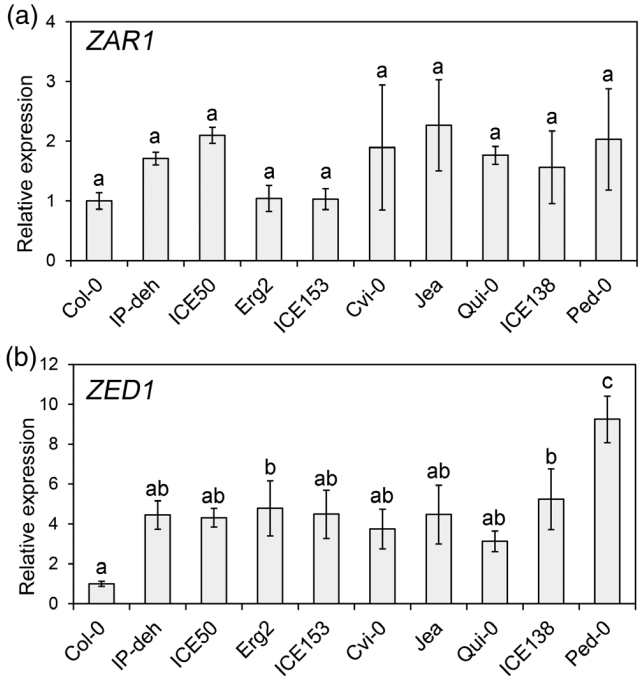



**Figure 2. *ZAR1* and *ZED1* are expressed in Arabidopsis ecotypes.** qRT-PCR analysis of gene expression in Arabidopsis ecotypes relative to Col-0. Relative gene expression were normalized to the endogenous *TIP41* transcript level [59] using $2^{-\Delta\Delta CT}$ method [60]. Error bars indicate the standard error from the mean of three independent biological replicates. The letters beside the bars indicate significance groups, as determined by a one-way ANOVA comparison followed by a Tukey's post hoc test ($P \leq 0.05$). A. *ZAR1* expressionB. *ZED1* expression.

*ZED1 ecotype variants are not impaired in HopZ1a recognition in N. benthamiana*

We previously showed that the co-expression of $ZED1^{Col-0}$ and HopZ1a in *N. benthamiana* leads to a strong HR dependent on the endogenous *NbZAR1* [34]. This HR is dependent on the expression of $AtZED1^{Col-0}$ as the expression of HopZ1a alone does not induce an HR, indicating that endogenous *N. benthamiana* kinases cannot recognize HopZ1a [34]. Weak mutations of ZED1 do not affect the HR when co-expressed with HopZ1a in *N. benthamiana*, while strong mutations like $ZED1^{D231N}$ abolish the HR [34]. To investigate the impact of the ZED1 polymorphisms on the HopZ1a-triggered HR, we co-expressed each *ZED1* allele identified in the ecotype screen with HopZ1a in *N. benthamiana* (Figure 1). All the *ZED1* alleles, including $ZED1^{Ped-0}$ carrying A305V, were able to induce a HR with similar levels of ion leakage as $ZED1^{Col-0}$. While it is possible that polymorphisms in ZED1 impair its function in *A. thaliana*, our *N. benthamiana* assay showed that they were all able to facilitate HopZ1a recognition. As well, we can conclude that none of the ZED1 polymorphisms strongly impacted ZED1 function,

as was observed with ZED1$^{D231N}$ (Figure S4). We thus focused on the effect of the ZAR1 polymorphisms for the remainder of our studies.

*Complementation of NbZAR1-VIGS by ZAR1 ecotype variants in N. benthamiana*

To investigate if the ZAR1 polymorphisms resulted in the loss of HopZ1a recognition, we carried out a complementation assay in *N. benthamiana* plants that were silenced for *NbZAR1*. The downregulation of *NbZAR1* by Virus-Induced Gene Silencing (VIGS) can be complemented by the expression of *AtZAR1*, and is visualized by the development of a HR and high levels of ion leakage when co-expressed with AtZED1$^{Col-0}$ and HopZ1a [34]. As a proof-of-principle, we tested *zar1* loss-of-function mutations previously identified in *A. thaliana* genetic screens [34,39,44] for their capacity to complement *NbZAR1*-VIGS. These previous screens identified five mutations that impair HopZ1a recognition (*zar1*$^{V202M}$, *zar1*$^{S291N}$, *zar1*$^{L465F}$, zar1$^{G645E}$ and *zar1*$^{S831F}$ [34,44]) and three mutations that impair AvrAC recognition (*zar1*$^{K195N}$, *zar1*$^{P359L}$ and *zar1*$^{P816Q}$ in Arabidopsis [62]) (Figure 3A). We therefore co-expressed each *zar1* mutant with HopZ1a and ZED1 in *NbZAR1*-silenced *N. benthamiana* leaves. To quantitively measure the efficiency of complementation, we monitored the strength of the HR by conductivity (Figure 3B). The *zar1*$^{K195N}$, *zar1*$^{P359L}$, *zar1*$^{L465F}$ and *zar1*$^{G645E}$ mutants were unable to complement *NbZAR1* with conductivity levels not statistically different from the negative control (EV). This strong phenotype was expected, since these mutations had previously been shown to be loss-of-function mutations in Arabidopsis [34,39]. The lack of complementation was not due to a lack of protein, as shown in Western blot analysis (Figure S5) [34]. Interestingly the *zar1*$^{S291N}$ and *zar1*$^{P816Q}$ mutants showed partial complementation that was significantly higher than EV. The *zar1*$^{V202M}$ and *zar1*$^{S831F}$ mutants showed even higher ion leakage levels that were statistically in the same group as the positive control ZAR1 (Figure 3B). These data suggest that K195, P359, L465 and G645 make strong contributions to ZAR1 function, compared to S291 and P816, and that V202 and S831 have particularly subtle effects. The data also show that the transient assay can distinguish among mutations of different strengths. *zar1*$^{V202M}$, *zar1*$^{S291N}$, *zar1*$^{P816Q}$ and *zar1*$^{S831F}$ all are loss-of-function mutations in Arabidopsis, but clearly still retain some residual activity which may be observed from over-expression in a heterologous system.

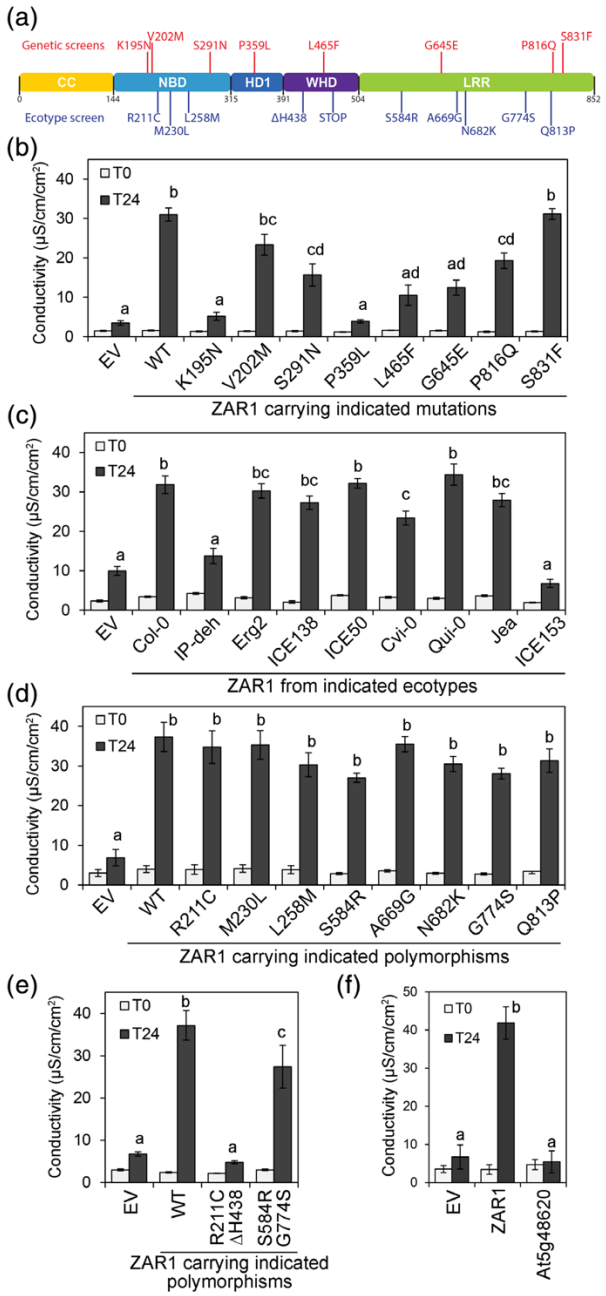



**Figure 3. Polymorphisms in ZAR1 can impact its ability to complement *NbZAR1*-VIGS.** A. ZAR1 sequence schematic showing the three main domains. The colored boxes correspond to the following domains: yellow is the CC domain, medium blue is the nucleotide-binding domain (NBD) subdomain, dark blue is the helical domain 1 (HD1) subdomain, purple is the winged helix domain (WHD) subdomain, and green is the leucine-rich repeat (LRR) domain. The amino acid changes induced by the mutation are indicated above the schematic for the genetic screens and below the schematic for the ecotype screen. B-F. Full-length ZAR1 with mutations exhibit partial or no complementation of *NbZAR1*–VIGS. Levels of ion leakage induced by ZED1$^{Col-0}$ and HopZ1a co-expressed with ZAR1, ZAR1 mutants, At5g48620, or EV were measured at 4 h (T0) and 28 h (T24) after dexamethasone application. Error bars indicate the standard error from the mean of six samples. The letters beside the bars indicate significance groups, as determined by a one-way analysis of variance (ANOVA) comparison followed by a Tukey's post hoc test ($P \leq 0.05$). The experiment was repeated three times with similar results. B. ZAR1 carrying mutations identified from genetic screens for loss of AvrAC and HopZ1a recognition [34,39,44] show diverse complementation of *NbZAR1*-VIGS. C. *ZAR1* alleles cloned from the IP-deh, Erg2, ICE138, ICE50, Cvi-0, Qui-0, Jea and ICE153

ecotypes show diverse complementation of *NbZAR1*-VIGS. D. Unique ZAR1 polymorphisms from the *ZAR1* allele in the IP-deh, Erg2, ICE138, ICE50, Cvi-0, Qui-0 and Jea ecotypes, were cloned into the Col-0 allele of *ZAR1* and were all able to complement *NbZAR1*-VIGS. E. Stacked unique ZAR1 polymorphisms from the IP-deh or Cvi-0 ecotypes cloned into the Col-0 allele of *ZAR1*, show the same level of complementation as the *ZAR1* allele from IP-deh or Cvi-0. F. ZAR1 from Col-0 complements *NbZAR1*-VIGS, while At5g48620 does not.

Since the complementation assay can distinguish subtle functional differences among ZAR1 alleles, we cloned *ZAR1* out of the nine accessions that lacked a HopZ1a-induced HR (Figure 1B). These *ZAR1* alleles carry specific polymorphisms that are only found in ETI-deficient groups. However, they carry additional substitutions that do not appear to contribute to recognition since these substitutions were found in *ZAR1* sequences from ETI-sufficient ecotypes. We observed a wide range of *NbZAR1*-VIGS complementation with *ZAR1* alleles from the ecotypes (hereafter ZAR1$^{\text{ecotype name}}$) co-expressed with ZED1$^{\text{Col-0}}$ and HopZ1a (Figure 3C). ZAR1$^{\text{IP-deh}}$ and ZAR1$^{\text{ICE153}}$ displayed a complete absence of complementation with ion leakage levels similar to EV. This phenotype suggests that the absence of HopZ1a-induced HR in these ecotypes is due to polymorphisms in the ZAR1 sequence. In fact, ZAR1$^{\text{ICE153}}$ carries a misannotated stop codon and lacks the end of NBARC (WHD) as well as the LRR domain, which validates our diversity screen. ZAR1$^{\text{Cvi-0}}$ displayed intermediate and statistically different complementation from ZAR1$^{\text{Col-0}}$ (Figure 3C). ZAR1$^{\text{ICE50}}$, ZAR1$^{\text{Erg2}}$, ZAR1$^{\text{Jea}}$, ZAR1$^{\text{Qui-0}}$ and ZAR1$^{\text{ICE138}}$ showed complete complementation with ion leakage levels similar to those of ZAR1$^{\text{Col-0}}$ and in the same statistical group (Figure 3C). All ZAR1 proteins were equally expressed in *N. benthamiana* except ZAR1$^{\text{ICE153}}$ which could not be detected due to an early stop codon that did not allow the translation of the C-terminal Myc tag (Figure S5). To investigate if the specific polymorphisms found in the ETI-deficient groups were sufficient to cause the complementation phenotype, we then introduced these substitutions into the *ZAR1*$^{\text{Col-0}}$ sequence and carried out the same complementation assay (Figure 3D). Interestingly, individual substitutions in *ZAR1*$^{\text{Col-0}}$ did not affect the ability of *ZAR1* to complement *NbZAR1*-VIGS (Figure 3D). However, *ZAR1*$^{\text{IP-deh}}$ and *ZAR1*$^{\text{Cvi-0}}$ carry two polymorphisms, that are specific to the ETI-deficient group (R211C/ΔH438 for ZAR1$^{\text{IP-deh}}$ and S584R/G774S for ZAR1$^{\text{Cvi-0}}$). To test if these polymorphisms might have an additive effect and impair ZAR1 function, we stacked the two corresponding polymorphisms in *ZAR1*$^{\text{Col-0}}$ and tested for complementation in *N. benthamiana*. Remarkably, the combination of polymorphisms restored the phenotype observed with ZAR1$^{\text{IP-deh}}$ or ZAR1$^{\text{Cvi-0}}$ (Figure 3E). The double mutant ZAR1$^{\text{R211C/ΔH438}}$ showed a similar complementation level as the negative control (EV) suggesting that the deletion of H438 combined with the R211C mutation strongly affects ZAR1 function (Figure 3D, 3E). The ZAR1$^{\text{S584R/G774S}}$ allele showed partial complementation similar to ZAR1$^{\text{Cvi-0}}$ with ion leakage levels significantly lower than ZAR1$^{\text{Col-0}}$ (Figure 3C and 3E). To show that this phenotype was specific for ZAR1, we tested whether a closely related RPP8-like NLR, At5g48620, could activate the HopZ1a-triggered HR. As expected, At5g48620 showed similar lack of complementation as EV (Figure 3F).

*Interaction between ZAR1 from ecotypes and ZED1*

We and others previously demonstrated that ZAR1 and ZED1 directly interact in the absence of pathogens [34,39,44]. This interaction is critical for the recognition of HopZ1a, as ZAR1 or ZED1 mutants impaired in this interaction are unable to induce the HR [34]. We thus carried out co-immunoprecipitation experiments (Co-IP) to test whether ZAR1 variants from the ecotypes could interact with ZED1$^{\text{Col-0}}$. We first cloned ZAR1 variants as in-frame fusions with a 5xMyc

epitope and ZED1$^{Col-0}$ as an in-frame fusion with a 3xFlag epitope. We co-expressed ZED1$^{Col-0}$-3xFlag with each ZAR1 variant in *N. benthamiana* leaves, followed by immunoprecipitation of protein complexes using anti-Myc magnetic beads. We also included At5g48620, a RPP8-like NLR, as a negative control. The nine ZAR1 variants cloned from the ecotypes retained the capacity to interact with ZED1$^{Col-0}$ as did the positive control ZAR1$^{Col-0}$. This interaction was specific since the negative control At5g48620 did not interact with ZED1 (Figure 4A). To investigate the impact of specific polymorphisms found in the ETI-deficient groups on the interaction with ZED1, we then introduced these substitutions into the *ZAR1*$^{Col-0}$ sequence and carried out the same Co-IP assay. ZAR1$^{Col-0}$ carrying the single substitutions or the double substitutions R211C/ΔH438 and S584R/G774S was able to interact with ZED1$^{Col-0}$ (Figure 4B), however ZAR1$^{Col-0}$ carrying R211C/ΔH438 interacted more weakly with ZED1$^{Col-0}$.

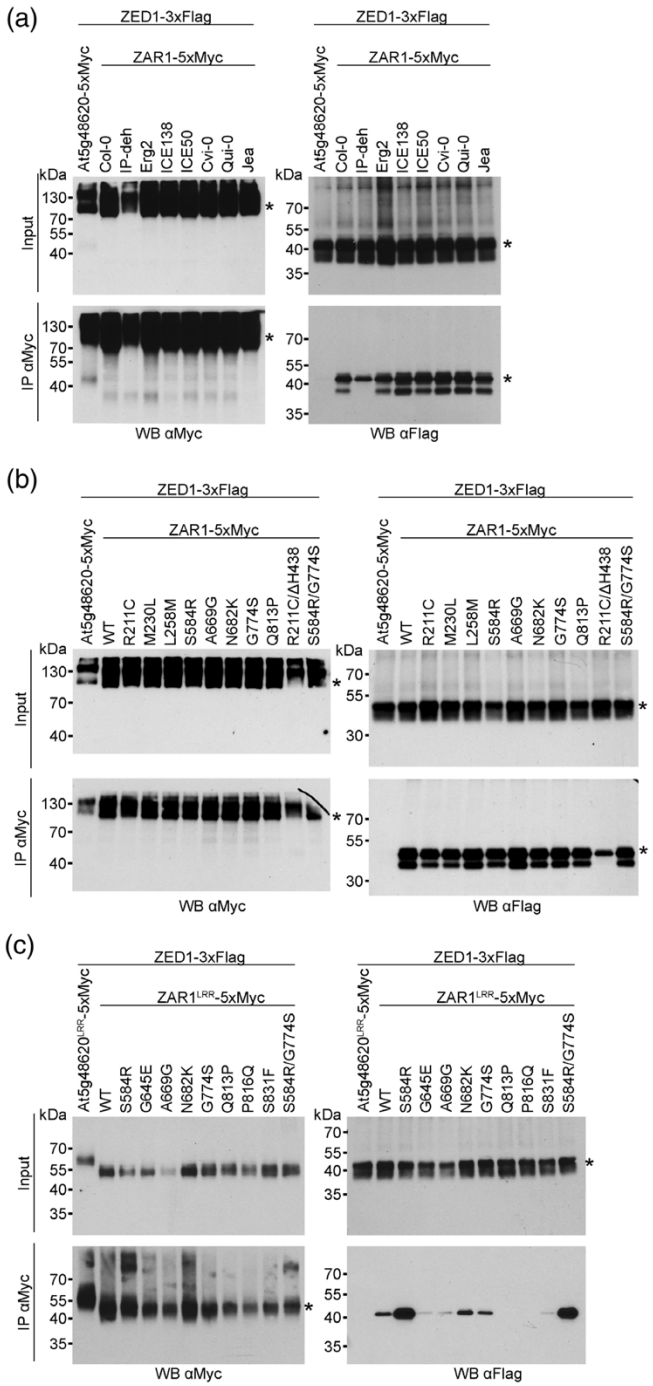



**Figure 4. ZAR1 polymorphisms S584R, G645E, Q813P, P816Q and S831F impact the association between ZAR1$^{LRR}$ and ZED1.** Immunoblots showing the co-immunoprecipitation of ZAR1 with ZED1 from *Nicotiana benthamiana* tissues. Western blot analysis (WB) was performed on the crude extract (input) or the immunoprecipitated fractions (IP) from anti-Myc beads. An asterisk indicates the band corresponding to the protein of interest, if multiple bands were observed. These experiments were repeated three times with similar results. A. At5g48620, *ZAR1* alleles from ecotypes with an in-frame 5xMyc epitope tag were co-expressed with ZED1 with an in-frame 3xFlag epitope tag in *N. benthamiana*. B. At5g48620, ZAR1 wild type (WT) or ZAR1 carrying polymorphisms with an in-frame 5xMyc epitope tag were co-expressed with ZED1 with an in-frame 3xFlag epitope tag in *N. benthamiana*. C. At5g48620$^{LRR}$, ZAR1$^{LRR}$ wild type (WT) or ZAR1$^{LRR}$ carrying polymorphisms fused to a 5xMyc epitope tag were co-expressed with ZED1 fused to a 3xFlag epitope tag in *N. benthamiana*.

We previously demonstrated that the CC domain of ZAR1 (amino acids 1-144) interacts with ZED1 in planta and in yeast [34,44], and that the LRR domain of ZAR1 (amino acids 505-852) is sufficient and required for Co-IP with ZED1 [34]. We hypothesized that the effect of polymorphisms in the LRR domain might be obscured by protein-protein interactions between ZED1 and other domains of ZAR1. We therefore specifically tested the four polymorphisms identified in the LRR domain of ZAR1$^{Cvi-0}$ (S584R/G774S), ZAR1$^{Qui-0}$ (N682K) and ZAR1$^{Jea}$ (A669G) for association with ZED1$^{Col-0}$ in the context of the LRR domain alone. The polymorphisms identified in ZAR1$^{Cvi-0}$ were tested independently, and also stacked in the same construct. For comparison, we included the three loss-of-function substitutions located in the LRR domain, that were identified during previous genetic screens (G645E, P816Q, S831F) [34,39]. As a negative control, we used the LRR domain of At3g48620 (amino acids 517-908). We were able to replicate the loss of interaction between ZAR1$^{P816Q}$ and ZED1$^{Col-0}$ that was previously described by Wang and colleagues [39] (Figure 4C). ZAR1$^{LRR}$ carrying either of the two mutations identified in the genetic screen (G645E or S831F) retained some interaction with ZED1$^{Col-0}$ but the interactions were significantly weaker than wild type ZAR1$^{LRR}$. This was consistent with our previous results using full length ZAR1 [34]. For the polymorphisms identified in the ecotype screen, ZAR1$^{LRR}$ carrying the S584R substitution appeared to stabilize the interaction with ZED1$^{Col-0}$, which was also observed in ZAR1$^{LRR}$ carrying the double substitution S584R/G774S found in Cvi-0 (Figure 4C). The A669G, N682K and G774S substitutions in ZAR1$^{LRR}$ did not impact the interaction with ZED1$^{Col-0}$, as the differences observed in the coimmunoprecipitation of ZED1$^{Col-0}$ appeared to correlate with the differences in the expression level of the ZAR1$^{LRR}$ constructs. Finally, ZAR1$^{LRR}$ carrying the Q813P substitution from ICE138 did not interact with ZED1$^{Col-0}$ (Figure 4C).

In conclusion, for full-length ZAR1, none of the polymorphisms strongly affected the interaction with ZED1$^{Col-0}$. However ZAR$^{Col-0}$ carrying the double polymorphism R211C/ΔH438 from IP-deh interacted more weakly with ZED1$^{Col-0}$. When we focused on the specific interaction between ZAR1$^{LRR}$ and ZED1, we found that ZED1$^{Col-0}$ interacted more strongly with ZAR1$^{LRR}$ carrying the S584R or S584R/G774S polymorphisms from Cvi-0 and more weakly with ZAR1$^{LRR}$ carrying a polymorphism from ICE138 (Q813P). These differences in interaction affinity could explain the weaker recognition phenotype observed with these accessions.

*HopZ1a recognition is impaired in Arabidopsis ecotypes*

Our results from the transient assay demonstrated that some polymorphisms had no effect on complementation of *NbZAR1*-VIGS or interaction with ZED1$^{Col-0}$. We therefore decided to re-confirm the phenotype of the nine ecotypes identified in the primary and secondary screens (IP-deh, Erg2, ICE138, ICE153, ICE50, Cvi-0, Qui-0, Jea and Ped-0), by repeating the HR assay on a greater number of plants infiltrated with *Pto*DC3000 (*ev*) and *Pto*DC3000 (*hopZ1a*). We phenotyped the HR 16 hours post-infiltration as no HR, weak HR, medium HR, or strong HR (Figure 5A, 5B). As previously observed, IP-deh, Cvi-0, ICE138, ICE153 and Jea lacked a HR to *Pto*DC3000 (*hopZ1a*), similar to the loss of HR observed with the knockout mutant *zar1-1* [38]. Ped-0 was mostly unable to induce a HopZ1a-triggered HR, and ~70% of the tested leaves showed a "no HR" phenotype. ICE50 had intermediate phenotypes with ~30% of plants showing no HR in response to *Pto*DC3000 (*hopZ1a*). However, the HRs in Erg2 and Qui-0 were most similar to those in Col-0 with respectively ~60% and ~80% of plants showing a strong HR when infiltrated with *Pto*DC3000 (*hopZ1a*) (Figure 5B). Taken together with our assays in *N. benthamiana*, these data indicate the polymorphisms found in Erg2 and Qui-0 (L258M or

N682K) have no effect on ZAR1 function. As well, these data confirm that IP-deh, Cvi-0, ICE138, ICE153, ICE50, Jea and Ped-0 are impaired in HopZ1a recognition.

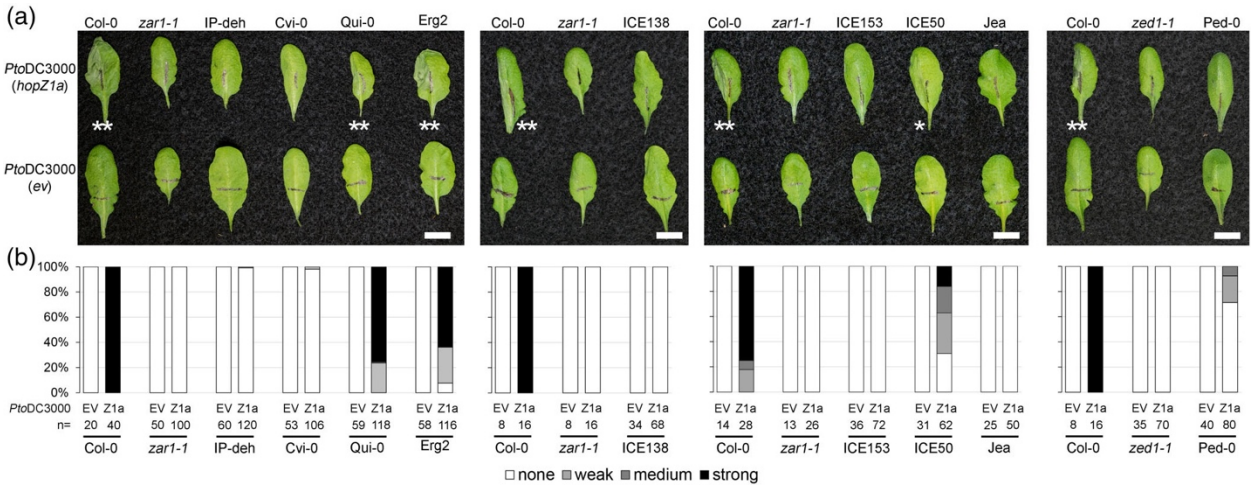



**Figure 5. HopZ1a recognition is lost in some Arabidopsis ecotypes.** Half-leaves of Arabidopsis Col-0, *zar1-1*, IP-deh, Cvi-0, Qui-0, Erg2, ICE138, ICE153, ICE50, ICE50, Jea, *zed1-1* or Ped-0 were infiltrated with *Pto*DC3000 carrying empty vector (*ev*) or *hopZ1a*. The bacteria were pressure-infiltrated into the leaves at $5 \times 10^7$ cfu/mL. A. Photos of representative phenotype were taken 20 h after infiltration. Strong HRs are marked with a double asterisk. Medium HRs are marked with a single asterisk. The data are from one experiment. Bars = 1 cm. B. For each ecotype, the percentage of leaves showing the HR phenotype 20 h after infiltration is shown after being scored on a scale from no HR, weak HR, medium HR to strong HR. "n" is the total number of leaves observed for each category.

*Stable transformation of ZAR1$^{Col-0}$ in IP-deh and Cvi-0 restores HopZ1a recognition*

Our Arabidopsis screen and complementation assay in *N. benthamiana* strongly suggested that the polymorphisms in ZAR1 from IP-deh and Cvi-0 particularly impaired ZAR1 function, and are thus responsible for the ETI-deficient phenotype observed in these lines. To test this hypothesis, we transformed IP-deh or Cvi-0 with a ZAR1$^{Col-0}$-5xMyc construct expressed under its native promoter. The first generation ($T_1$) was selected for resistance to the herbicide glufosinate ammonium (Basta) and the presence of the transgene was confirmed by PCR. We tested for restoration of the immune response by infiltrating the plants with *Pto*DC3000 (*ev*) and *Pto*DC3000 (*hopZ1a*). The four complemented IP-deh lines showed medium to strong HRs in response to *Pto*DC3000 (*hopZ1a*) similar to Col-0, while *zar1-1* and IP-deh wild type plants showed no recognition of HopZ1a (Figure 6A). The two complemented Cvi-0 lines and the two complemented *zar1* lines showed strong HRs in response to *Pto*DC3000 (*hopZ1a*) similar to Col-0, while *zar1-1* and Cvi-0 wild type plants were unable to recognize HopZ1a (Figure 6A). For each line, the expression of the construct was tested by western blot analysis (Figure 6B). We observed a correlation between the level of expression of the ZAR1$^{Col-0}$-Myc construct and the level of complementation (Figure 6). This validates that the polymorphisms in ZAR1$^{IP-deh}$ and ZAR1$^{Cvi-0}$ are responsible for the ETI-deficient phenotype of IP-deh and Cvi-0.

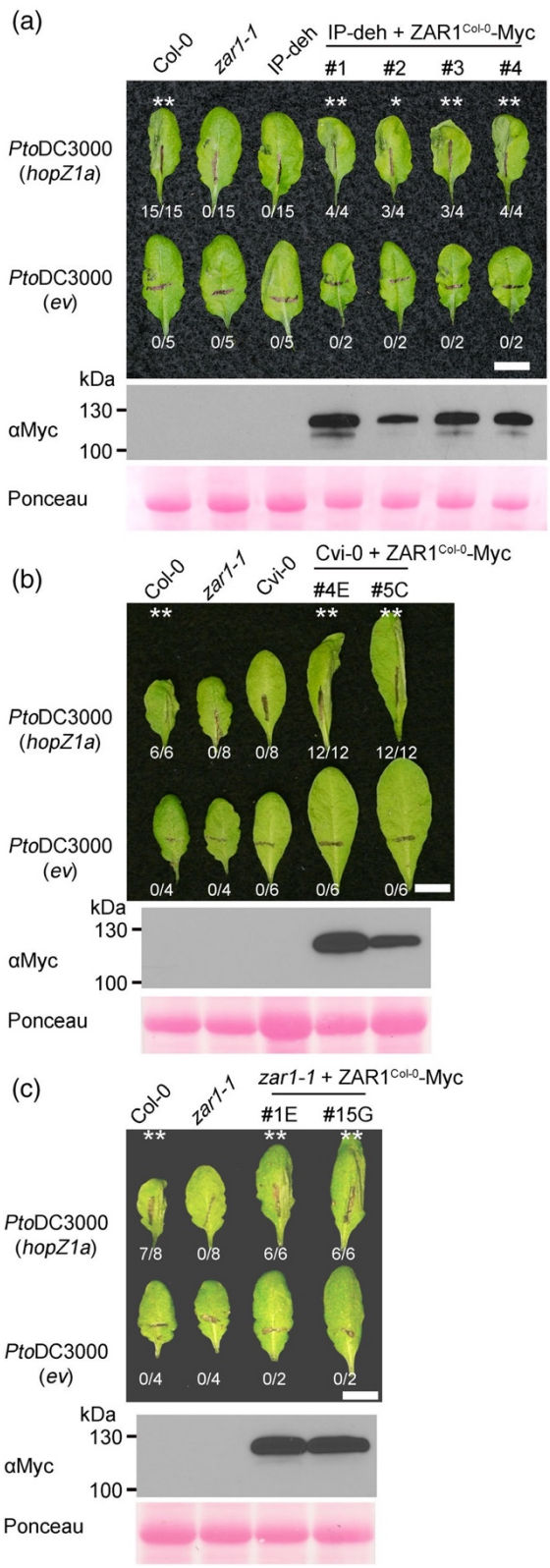



**Figure 6. ZAR1$^{Col-0}$ complements ETI-deficiency observed in IP-deh, Cvi-0 and *zar1-1*.** Half-leaves of Arabidopsis plants were infiltrated with *Pto*DC3000 carrying empty vector (*ev*) or *hopZ1a*. IP-deh plants were T1 transgenic lines, and Cvi-0 or *zar1-1* plants were T3 transgenic lines. The bacteria were pressure-infiltrated into the

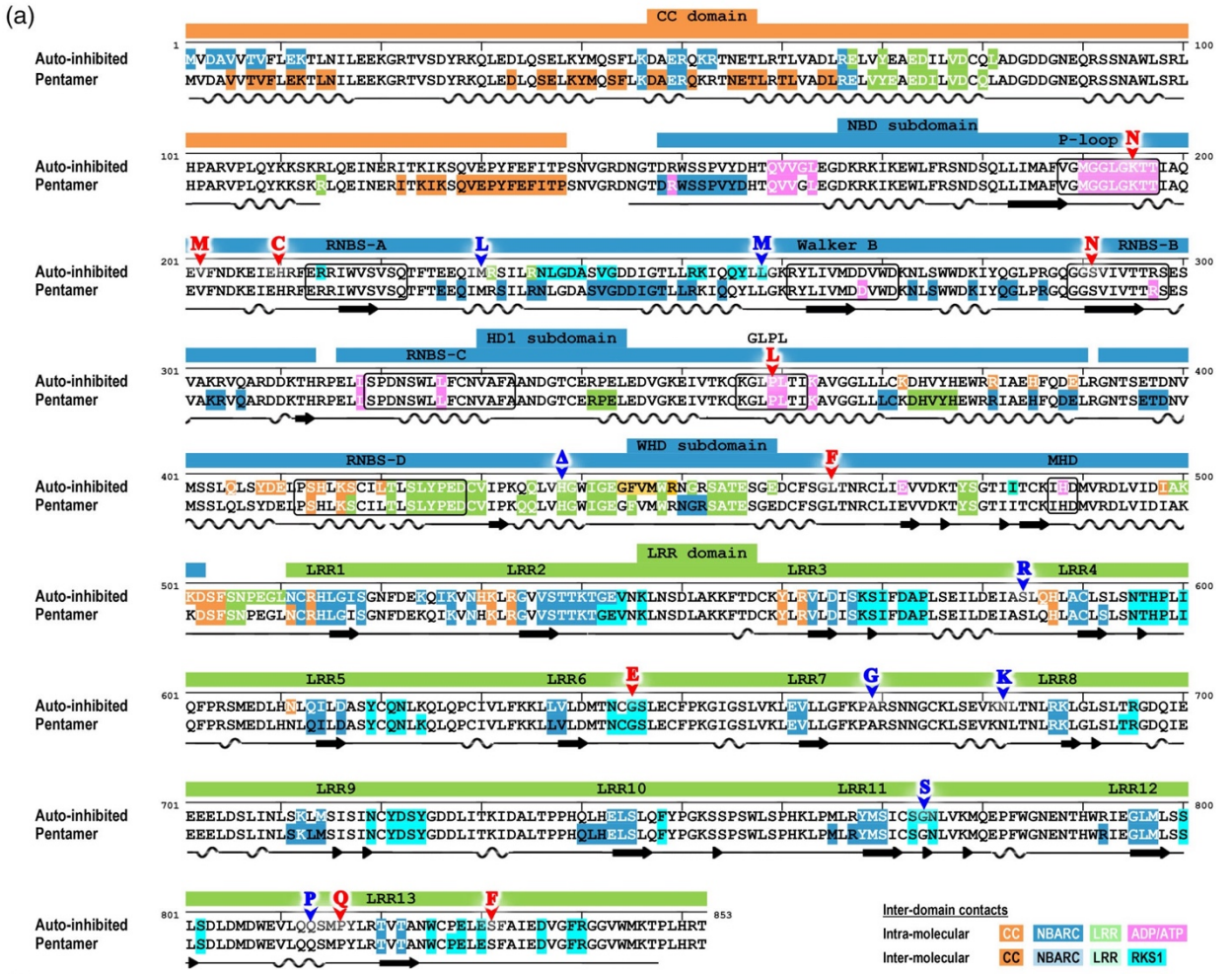
leaves at $5 \times 10^7$ cfu/mL. Photos of representative leaves were taken 20 h after infiltration. Strong HRs are marked with a double asterisk. Medium HRs are marked with a single asterisk. HRs on T1 plants were performed once, and HRs on T3 plants were performed twice. Bars = 1 cm. ZAR1 protein expression was detected with an α-Myc antibody, from a single leaf of 6 week-old Arabidopsis plants. Ponceau staining of total protein extract was used as a loading control. A. Arabidopsis Col-0, *zar1-1*, IP-deh or IP-deh transgenic lines carrying $pZAR1^{Col\text{-}0}::ZAR1^{Col\text{-}0}:5xMyc$. B.Arabidopsis Col-0, *zar1-1*, Cvi-0 or Cvi-0 transgenic lines carrying $pZAR1^{Col\text{-}0}::ZAR1^{Col\text{-}0}:5xMyc$. C. Arabidopsis Col-0, *zar1-1*, or *zar1-1* transgenic lines carrying $pZAR1^{Col\text{-}0}::ZAR1^{Col\text{-}0}:5xMyc$.

## Discussion

ZAR1 is central to an ancient and conserved immune network that recognizes multiple bacterial effectors [35,36,38–40]. Here we demonstrate that natural diversity in association with structure-function analyses can be powerful tools to identify essential determinants of immune receptors. Through a targeted screen of the 1001 genomes collection (Figure 1), we identified seven ecotypes that lack ETI and carry unique polymorphisms in ZAR1 or ZED1 (Figure 2, 5). The polymorphism from Ped-0 in ZED1 is located at the same position we previously identified in a forward genetics screen for the loss of the HopZ1a induced HR (Figure 1) [34,44]. We identified one polymorphism in ICE138 that results in weaker interactions between $ZAR1^{LRR}$ and ZED1, which might account for the loss of HR in Arabidopsis (Figure 4, 5), as well as an misannotated stop codon in ICE153 that prematurely truncates ZAR1 and removes the ZED1 interaction domain (Figure 2, 5) [34]. For Cvi-0 and IP-deh, we identified one pair of polymorphisms in ZAR1 that results in reduced immune responses (Figure 3, 5), and affects interactions between ZAR1 and ZED1 (Figure 4). Complementation of IP-deh or Cvi-0 with the Col-0 allele of *ZAR1* restores the immune response to *P. syringae* carrying HopZ1a (Figure 6), indicating that specific polymorphisms in ZAR1 affect its function in IP-deh or Cvi-0.

We previously carried out a detailed structure-function analysis of ZAR1 [33] that took advantage of the cryo-EM structures of ZAR1 in association with RKS1 in the inactive and active states [29,48], to understand the role of specific residues in intra- and intermolecular interactions. We extended these analyses to evaluate the impact of the naturally occurring polymorphisms identified in this study as well as the mutations identified in previous genetic screens [34,39] on ZAR1 structures (Figure 7, S6). Two mutations in the NBARC (NBD/HD1) domain, K195N and P359L (Wang et al., 2015), are involved in nucleotide binding, and therefore have a very strong impact on ZAR1 function as the mutants were not able to complement *NbZAR1*-VIGS. The deletion of H438 has strong effect on ZAR1 function, and might disrupt the folding of ZAR1, as the cryo-EM structures showed that this residue is involved in the interaction between the WHD subdomain of the NBARC domain, and LRR domain in the inactive state. While R211C had no effect on complementation of *NbZAR1* silencing (Figure 3D) or the interaction with ZAR1 (Figure 4B), we cannot exclude the possibility that the combination of R211C/ΔH438 from ZAR1 in IP-deh affects ZAR1 function. The remaining polymorphisms (V202M, R211C, M230L, and L465F) identified in the NBARC domain are not located in obvious motifs or involved in intra- or intermolecular contacts predicted by the cryo-EM structures of ZAR1 (Figure 7B). ZAR1 carrying L465F was not able to complement *NbZAR1*-VIGS silenced plants, indicating that this residue is functionally important (Figure 3).

(a)




(b)

| Screen | Residue | Position | Mutation | Blosum62 score | Domain | Motif | RSA | Complementation in *N. benthamiana* | ZAR1$^{LRR}$-ZED1 interaction | Interaction in inactive state (6J5W) | Interaction in activated pentameric state (6J5T) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AvrAC | Lys | 195 | Asn | 0 | NBD | P-loop | B | no | N/A | ADP | ATP |
| HopZ1a | Val | 202 | Met | 1 | NBD | | B | med | N/A | - | |
| Ecotype | Arg | 211 | Cys | -3 | NBD | | E | yes | N/A | - | |
| Ecotype | Met | 230 | Asn | -2 | NBD | | B | yes | N/A | - | |
| Ecotype | Leu | 258 | Met | 2 | NBD | | E | yes | N/A | RKS1 | - |
| HopZ1a | Ser | 291 | Asn | 1 | NBD | RNBS-B | B | med | N/A | - | - |
| AvrAC | Pro | 359 | Leu | -3 | HD1 | GLPL | B | no | N/A | ADP | ATP |
| Ecotype | His | 438 | - | - | WHD | | B | no | N/A | LRR intra-molecular | LRR intra-molecular |
| HopZ1a | Leu | 465 | Phe | 0 | WHD | | B | med | N/A | - | - |
| Ecotype | Ser | 584 | Arg | -1 | LRR | outer side | E | yes | strong | - | - |
| HopZ1a | Gly | 645 | Glu | -2 | LRR | upper side | E | med | weak | RKS1 | RKS1 |
| Ecotype | Ala | 669 | Gly | 0 | LRR | upper side | B | yes | weak | - | - |
| Ecotype | Asn | 682 | Lys | 0 | LRR | outer side | E | yes | yes | - | - |
| Ecotype | Gly | 774 | Ser | 0 | LRR | upper side | E | yes | yes | RKS1 | - |
| Ecotype | Gln | 813 | Pro | -1 | LRR | outer side | E | yes | no | - | - |
| AvrAC | Pro | 816 | Gln | -1 | LRR | outer side | E | med | no | - | - |
| HopZ1a | Ser | 831 | Phe | -2 | LRR | outer side | E | yes | weak | - | - |

**Figure 7. Structure-informed intra- and intermolecular contacts in ZAR1.** A. ZAR1 interdomain contacts in both the auto-inhibited ADP-bound ZAR1-RKS1 conformation (PDB: 6j5w) and the dATP-bound pentameric conformation (PDB: 6j5t) from the cryo-EM structures. In the inactive monomeric state, only intramolecular ZAR1 interactions and intermolecular ZAR1-RKS1 interactions are present. In the pentameric activated state, additional intermolecular ZAR1-ZAR1 interactions are shown in black font. Domains are colored as follows: CC – orange; NBARC - blue, LRR – green; RKS1 – cyan. Interdomain contacts are mapped onto the sequence using a color code corresponding to the interacting domain partner. Contacts within 5Å are shown. In cases where contacts below this threshold are formed with multiple domains, only the most substantial contact pair is represented. The secondary structure in the auto-inhibited conformation (PDB: 6j5w) is shown below the sequence as follows: helices (wavy lines), beta sheets or strands (arrows), other structure (straight lines). Unsolved coordinates do not have a secondary structure assignment. Polymorphisms from the ecotype (blue) and genetic (red) screen are indicated with arrows. B.

Summary of data for each polymorphism studied in this manuscript. For each mutation, the Blosum62 and relative solvent accessibility (RSA) are indicated. Complementation data are compiled from Figure 3, ZAR1$^{LRR}$-ZED1 interactions are from Figure 4C and structural data are from Figure 7A.

From the ecotype screen and the forward genetics screen, eight mutations or polymorphisms were identified in the LRR domain. According to the cryo-EM structures, only G645 and G774 are involved in the interaction with RKS1 (Figure 7, S6). When we specifically evaluated the impact of these polymorphisms on the ZAR1$^{LRR}$-ZED1 interaction (Figure 4C), ZAR1$^{LRR}$ carrying G645E interacted more weakly with ZED1, while ZAR1$^{LRR}$ carrying G774S interacted with ZED1 to the same extent as ZAR1$^{LRR}$. We confirmed that the P816Q mutation abolished the ZED1 interaction [34,39] and we also showed that Q813P (from ICE138) had a similar impact. Finally, ZAR1$^{LRR}$ carrying the S831F mutation displayed reduced affinity for ZED1, while the S584R substitution seemed to stabilize this association. Interestingly, none of these four residues (S584, Q813, P816 and S831) are predicted to play a direct role in the RKS1 interaction according to the cryo-EM structures (Figure 7, S6). We hypothesize that these mutations/polymorphisms have a more systemic effect on the ZAR1$^{LRR}$ conformation or that the ZAR1-ZED1 interaction requires slightly different residues than the ZAR1-RKS1 interaction. The enhanced interaction of ZED1 with ZAR1 carrying S584R or S584R/G774S might prevent conformational changes in the LRR domain after HopZ1a recognition (Figure 4C). Most of the eight polymorphisms in the LRR domain were still able to complement, at least partially, *NbZAR1*-VIGS (Figure 3). Therefore, these residues are not essential for the transduction of HopZ1a-induced immunity in *N. benthamiana.* Alternatively, a weak allele may function in *N. benthamiana* if it is overexpressed but be non-functional under native expression in Arabidopsis. Our work also highlights the importance of analyzing the effect of different mutations in both Arabidopsis and *N. benthamiana*. We did not identify any polymorphisms in the CC domain of ZAR1, which is consistent with our previous work showing that the CC domain is very sensitive to structural changes [33].

Natural diversity observed in immune receptors is often the result of local adaptation to effector variability and because of the arms race between plants and pathogens. This has been particularly demonstrated for NLRs that directly interact with their cognate effector proteins. Different alleles of the NLR RPP1 [RECOGNITION OF PERONOSPORA PARASITICA 1] recognize variants of the *Hyaloperonospora arabidopsidis* effector ATR1 [63–65]. Similarly, allelic diversification of the MLA [MILDEW RESISTANCE LOCUS A] immune receptor allows for the recognition of multiple variants of $AVR_A$, an effector protein from *Blumeria graminis* f. sp. *hordei* [66]. The flax L7 NLR binds more weakly to AvrL567, an effector from the fungal pathogen *Melampsora lini*, compared to the L6 NLR, due to polymorphisms in the TIR and NB domains [67]. Diversity has also been observed across plant species in the guarded targets of NLRs, such as RIN4 [RPM1 INTERACTING PROTEIN 4] [68]. RIN4 is targeted by several unrelated effector proteins from *P. syringae*, including AvrB and AvrRpm1, as well as AvrRpt2 from *P. syringae* and *Erwinia amylovora* [69–72]. NLRs have evolved independently in multiple plant species to recognize modifications of RIN4, and polymorphisms in RIN4 are critical for the activation for these different NLRs [73–80]. In contrast, ZAR1 is an ancient NLR, which has retained its function through different plant lineages [34,36,38,42]. The interaction between ZAR1 and effector proteins is mediated by ZRKs [41]. Since ZAR1 interacts with ZED1, RKS1/ZRK1, ZRK3, ZRK6 and ZRK15 [39,40,44], we speculate that natural variability, particularly in ZAR1$^{LRR}$ domain can shift the ZAR1/ZRK equilibrium towards the recognition of frequently encountered effectors. Neofunctionalization of the ZRKs may have played a greater role in expanding the range of

effectors recognized by ZAR1. Consistent with this, our natural diversity screen identified polymorphisms in the LRR domain that impact the interaction between ZAR1 and ZED1 (Figure 4C). A thorough analysis of ZAR1-ZRK interactions might reveal ZAR1$^{LRR}$ residues that provide interaction specificity of ZAR1 with ZRKs.

Our work illustrates how natural diversity can be mined to understand functional effects on immunity and NLR activation. ZAR1 is likely able to recognize many different effector proteins through diversity in its LRR domain as well as the specific RLCK it interacts with. Our screen also identified ecotypes that were unable to trigger HopZ1a immunity despite having similar ZAR1 and ZED1 protein sequences as ETI-sufficient ecotypes. We cannot exclude the possibility that another effector in *P. syringae* suppresses HopZ1a recognition in these ecotypes. However it is also possible that these ecotypes are deficient in an unknown component of the ZAR1 immune pathway. Two recent studies suggest that other kinases in addition to ZED1 also play a role in HopZ1a-mediated immunity. Two kinases, SZE1 and SZE2 [SUPPRESSOR OF ZED1-D1], have been shown to redundantly participate in HopZ1a-induced immunity [47]. SZE1 and SZE2 directly interact with HopZ1a and are proposed to help mediate conformational changes that lead to ZAR1 activation [47]. In addition, several PBS1-like kinases (PBLs) have also been shown to be targeted by HopZ1a and participate in ZAR1-mediated recognition [81]. It would therefore be interesting to broaden our diversity screen to these other players in HopZ1a-induced immunity and investigate how evolution shaped these proteins in relation to ZED1 and ZAR1. Our work shows that natural diversity can help to reveal parts of immune receptors that are more plastic or more constrained by evolution. This information will be crucial for rational engineering of immune receptors.

**References**


1. Huard-Chauveau, C. *et al.* An Atypical Kinase under Balancing Selection Confers Broad-Spectrum Disease Resistance in Arabidopsis. *PLoS Genet.* **9**, e1003766 (2013).
2. Glander, S. *et al.* Assortment of flowering time and immunity alleles in natural Arabidopsis thaliana populations suggests immunity and vegetative lifespan strategies coevolve. *Genome Biol. Evol.* **10**, 2278–2291 (2018).
3. Aoun, N. *et al.* Quantitative Disease Resistance under Elevated Temperature: Genetic Basis of New Resistance Mechanisms to Ralstonia solanacearum. *Front. Plant Sci.* **8**, 1–16 (2017).
4. Montes, N., Alonso-Blanco, C. & García-Arenal, F. *Cucumber mosaic virus* infection as a potential selective pressure on *Arabidopsis thaliana* populations. *PLOS Pathog.* **15**, e1007810 (2019).
5. Barragan, C. A. *et al.* RPW8/HR repeats control NLR activation in *Arabidopsis thaliana*. *PLoS Genet.* **15**, 1–21 (2019).
6. Van de Weyer, A.-L. *et al.* A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis thaliana. *Cell* **178**, 1260–1272 (2019).
7. Aranzana, M. J. *et al.* Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genet.* **1**, e60 (2005).
8. Mitchell-Olds, T. & Schmitt, J. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature* **441**, 947–952 (2006).
9. Weigel, D. & Mott, R. The 1001 Genomes Project for *Arabidopsis thaliana*. *Genome Biol.* **10**, 107 (2009).

10. Alonso-Blanco, C. *et al.* 1,135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491 (2016).
11. Cook, D. E., Mesarich, C. H. & Thomma, B. P. H. J. Understanding Plant Immunity as a Surveillance System to Detect Invasion. *Annu. Rev. Phytopathol.* **53**, 541–563 (2015).
12. Jones, J. D. G., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, 1117 (2016).
13. Lolle, S., Stevens, D. & Coaker, G. ScienceDirect Plant NLR-triggered immunity : from receptor activation to downstream signaling. *Curr. Opin. Immunol.* **62**, 99–105.
14. Boutrot, F. & Zipfel, C. Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annu. Rev. Phytopathol.* **55**, 257–286 (2017).
15. Saijo, Y., Loo, E. P. & Yasuda, S. Pattern recognition receptors and signaling in plant-microbe interactions. *Plant J.* **93**, 592–613 (2018).
16. Cesari, S. Multiple strategies for pathogen perception by plant immune receptors. *New Phytol.* **219**, 17–24 (2018).
17. Schreiber, K. J., Baudin, M., Hassan, J. A. & Lewis, J. D. Die another day: Molecular mechanisms of effector-triggered immunity elicited by type III secreted effector proteins. *Semin. Cell Dev. Biol.* **56**, 124–133 (2016).
18. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
19. Heath, M. C. Hypersensitive response-related death. *Plant Mol. Biol.* **44**, 321–334 (2000).
20. Khan, M., Seto, D., Subramaniam, R. & Desveaux, D. Oh, the places they'll go! A survey of phytopathogen effectors and their host targets. *Plant J.* **12**, 3218–3221 (2017).
21. Schreiber, K. J., Bentham, A., Williams, S. J., Kobe, B. & Staskawicz, B. J. Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLOS Pathog.* **12**, e1005769 (2016).
22. Dodds, P. N. & Rathjen, J. P. Plant immunity: towards an integrated view of plant–pathogen interactions. *Nat. Rev. Genet.* **11**, 539–548 (2010).
23. van der Hoorn, R. A. L. & Kamoun, S. From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**, 2009–2017 (2008).
24. Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T. & Dodds, P. N. A novel conserved mechanism for plant NLR protein pairs: The 'integrated decoy' hypothesis. *Front. Plant Sci.* **5**, 606 (2014).
25. Wu, C., Krasileva, K. V, Banfield, M. J., Terauchi, R. & Kamoun, S. The "sensor domains" of plant NLR proteins: more than decoys? *Front. Plant Sci.* **6**, 134 (2015).
26. Baggs, E., Dagdas, G. & Krasileva, K. V. K. V. NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Curr. Opin. Plant Biol.* **38**, 59–67 (2017).
27. Takken, F. L. W. & Goverse, A. How to build a pathogen detector: Structural basis of NB-LRR function. *Curr. Opin. Plant Biol.* **15**, 375–384 (2012).
28. Burdett, H. *et al.* The Plant "Resistosome": Structural Insights into Immune Signaling. *Cell Host Microbe* **26**, 193–201 (2019).
29. Wang, J. *et al.* Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science* **364**, (2019).
30. Collier, S. M. & Moffett, P. NB-LRRs work a 'bait and switch' on pathogens. *Trends Plant Sci.* **14**, 521–529 (2009).
31. 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* **9**, 200–211 (2011).
32. Maekawa, T. *et al.* Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* **9**, 187–199 (2011).
33. Baudin, M., Schreiber, K. J., Martin, E. C., Petrescu, A. J. & Lewis, J. D. Structure–function analysis of ZAR1 immune receptor reveals key molecular interactions for activity. *Plant J.* **101**, 352–370 (2019).
34. Baudin, M., Hassan, J. A., Schreiber, K. J. & Lewis, J. D. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol.* **174**, 2038–2053 (2017).
35. Laflamme, B. *et al.* The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. *Science* **367**, 763–768 (2020).
36. Schultink, A., Qi, T., Bally, J. & Staskawicz, B. Using forward genetics in Nicotiana benthamiana to uncover the immune signaling pathway mediating recognition of the Xanthomonas perforans effector XopJ4. *New Phytol.* **221**, 1001–1009 (2019).
37. Lewis, J. D., Abada, W., Ma, W., Guttman, D. S. & Desveaux, D. The HopZ family of Pseudomonas syringae type III effectors require myristoylation for virulence and avirulence functions in Arabidopsis thaliana. *J. Bacteriol.* **190**, 2880–2891 (2008).
38. Lewis, J. D., Wu, R., Guttman, D. S. & Desveaux, D. Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. *PLoS Genet.* **6**, e1000894 (2010).
39. Wang, G. *et al.* The decoy substrate of a pathogen effector and a pseudokinase specify pathogen-induced modified-self recognition and immunity in plants. *Cell Host Microbe* **18**, 285–295 (2015).
40. Seto, D. *et al.* Expanded type III effector recognition by the ZAR1 NLR protein using ZED1-related kinases. *Nat. Plants* **3**, 25–28 (2017).
41. Lewis, J. D., Lo, T., Bastedo, P., Guttman, D. S. & Desveaux, D. The rise of the undead: Pseudokinases as mediators of effector-triggered immunity. *Plant Signal. Behav.* **9**, e27563 (2014).
42. Ma, W., Dong, F. F. T., Stavrinides, J. & Guttman, D. S. Type III Effector Diversification via Both Pathoadaptation and Horizontal Transfer in Response to a Coevolutionary Arms Race. *PLoS Genet.* **2**, e209 (2006).
43. Lewis, J. D. *et al.* The YopJ superfamily in plant-associated bacteria. *Mol. Plant Pathol.* **12**, 928–937 (2011).
44. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the *Pseudomonas syringae*type III effector HopZ1a. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18722–18727 (2013).
45. Lee, A. H. Y. *et al.* A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion. *PLoS Pathog.* **8**, e1002523 (2012).
46. Martel, A. *et al.* Immunodiversity of the Arabidopsis ZAR1 NLR Is conveyed by receptor-like cytoplasmic kinase sensors. *Front. Plant Sci.* **11**, 1290 (2020).
47. Liu, C. *et al.* Two Arabidopsis Receptor-like Cytoplasmic Kinases SZE1 and SZE2 Associate with the ZAR1-ZED1 Complex and Are Required for Effector-Triggered Immunity. *Mol. Plant* **12**, 967–983 (2019).
48. Wang, J. *et al.* Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science* **364**, (2019).

49. Hu, M., Qi, J., Bi, G. & Zhou, J.-M. Bacterial effectors induce oligomerization of immune receptor ZAR1 in vivo. *Mol. Plant* **13**, 793–801 (2020).
50. Aoyama, T. & Chua, N.-H. A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.* **11**, 605–612 (1997).
51. Earley, K. W. *et al.* Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616–629 (2006).
52. Hayward, A., Padmanabhan, M. & Dinesh-Kumar, S. P. Virus-induced gene silencing in Nicotiana benthamianaand other plant species. *Methods Mol. Biol.* **678**, 55–63 (2011).
53. Clough, S. J. & Bent, A. F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
54. Katagiri, F., Thilmony, R. & He, S. Y. The Arabidopsis Thaliana-Pseudomonas Syringae Interaction. *Arab. B.* **1**, e0039 (2002).
55. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2018).
56. Ronquist, F. *et al.* Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
57. Nordborg, M. *et al.* The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol.* **3**, 1289–1299 (2005).
58. Pisupati, R. *et al.* Verification of Arabidopsis stock collections using SNPmatch, a tool for genotyping high-plexed samples. *Sci. Data* **4**, 170184 (2017).
59. Piisilä, M. *et al.* The F-box protein MAX2 contributes to resistance to bacterial phytopathogens in *Arabidopsis thaliana*. *BMC Plant Biol.* **15**, 1–17 (2015).
60. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* **25**, 402–408 (2001).
61. Ye, J. *et al.* Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 134 (2012).
62. Wang, G.-F. *et al.* Molecular and functional analyses of a maize autoactive NB-LRR protein identify precise structural requirements for activity. *PLOS Pathog.* **11**, e1004674 (2015).
63. Steinbrenner, A. D., Goritschnig, S. & Staskawicz, B. J. Recognition and Activation Domains Contribute to Allele-Specific Responses of an Arabidopsis NLR Receptor to an Oomycete Effector Protein. *PLoS Pathog.* **11**, 1–19 (2015).
64. Fabro, G. *et al.* Multiple candidate effectors from the oomycete pathogen Hyaloperonospora arabidopsidis suppress host plant immunity. *PLoS Pathog.* **7**, e1002348 (2011).
65. Botella, M. A. *et al.* Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct *Peronospora parasitica*avirulence determinants. *Plant Cell* **10**, 1847–1860 (1998).
66. Saur, I. M. *et al.* Multiple pairs of allelic MLA immune receptor-powdery mildew AVRA effectors argue for a direct recognition mechanism. *Elife* **8**, 1–31 (2019).
67. Bernoux, M. *et al.* Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *Plant Cell* **28**, 146–159 (2016).
68. Toruño, T. Y., Shen, M., Coaker, G. & Mackey, D. Regulated disorder: Posttranslational modifications control the RIN4 plant immune signaling hub. *Mol. Plant-Microbe Interact.* **32**, 56–64 (2019).

69. Mackey, D., Holt, B. F., Wiig, A. & Dangl, J. L. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743–754 (2002).
70. Mackey, D., Belkhadir, Y., Alonso, J. M., Ecker, J. R. & Dangl, J. L. Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* **112**, 379–389 (2003).
71. Axtell, M. J. & Staskawicz, B. J. Initiation of RPS2-specified disease resistance in Arabidopsis Is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**, 369–377 (2003).
72. Zhao, Y., He, S.-Y. & Sundin, G. W. The Erwinia amylovora avrRpt2EA gene contributes to virulence on pear and AvrRpt2EA is recognized by Arabidopsis RPS2 when expressed in Pseudomonas syringae. *Mol. plant-microbe Interact.* **19**, 644–654 (2006).
73. Prokchorchik, M. *et al.* A host target of a bacterial cysteine protease virulence effector plays a key role in convergent evolution of plant innate immune system receptors. *New Phytol.* **225**, 1327–1342 (2020).
74. Selote, D. & Kachroo, A. RPG1-B-derived resistance to AvrB-expressing Pseudomonas syringae requires RIN4-like proteins in soybean. *Plant Physiol.* **153**, 1199–1211 (2010).
75. Ashfield, T., Keen, N. T., Buzzell, R. I. & Innes, R. W. Soybean resistance genes specific for different Pseudomonas syringaea virulence genes are allelic, or closely linked, at the RPG1 locus. *Genetics* **141**, 1597–1604 (1995).
76. Ashfield, T. *et al.* Evolutionary Relationship of Disease Resistance Genes in Soybean and Arabidopsis Specific for the Pseudomonas syringae Effectors AvrB and AvrRpm1. *Plant Physiol.* **166**, 235–251 (2014).
77. Kessens, R., Ashfield, T., Kim, S. H. & Innes, R. W. Determining the GmRIN4 Requirements of the Soybean Disease Resistance Proteins Rpg1b and Rpg1r Using a *Nicotiana glutinosa*-Based Agroinfiltration System. *PLoS One* **9**, e108159 (2014).
78. Mindrinos, M., Katagiri, F., Yu, G.-L. & Ausubel, F. M. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78**, 1089–1099 (1994).
79. Bent, A. F. *et al.* RPS2 of Arabidopsis thaliana: a leucine-rich repeat class of plant disease resistance genes. *Science* **265**, 1856–1860 (1994).
80. Grant, M. R. *et al.* Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. *Science* **269**, 843–846 (1995).
81. Bastedo, D. P. *et al.* Perturbations of the ZED1 pseudokinase activate plant immunity. *PLoS Pathog.* **15**, e1007900 (2019).

Table S1: Oligonucleotides used in this study.

| Application | Primer Name | Sequence (5' to 3') |
|---|---|---|
| Cloning ZAR1 from Col-0 | 13-5'ZAR1ATGXhoI | GCGCTCGAGATGGTGGACGCTGTTGTAACAGT |
| | 14-3'ZAR1StuI | CTTCCGGAGGTTCTGTGCAATGGTGTTTTCAT |
| Cloning ZED1 from Col-0 | 9-5'ZED1XhoIATG | GCCCTCGAGATGAGCAAGAACAATAAGAAGAAGAGA |
| | 64-3'ZED1TEV | gaaagtacaagttttcAGAGAGTTTCTCAATCAATTTGAGC |
| | 109-3'FlagTGASpeI | CCACTAGTTCATTTGTCATCATCATCCTTGTAATCTTT |
| | 231-5'At3g57750bp265 | GCGTATAGAGATATCTCTGTATCCT |
| | 65-5'TEVZED1 | gattgagaaactctctGAAAACTTGTACTTTCAAGGCGAC |
| | 63-3'FlagattB2 | gaaagctgggtgTTTGTCATCATCATCCTTGTAATCTT |
| Cloning ZAR1 from ecotype | 13-5'ZAR1ATGXhoI | GCGCTCGAGATGGTGGACGCTGTTGTAACAGT |
| | 58-3'ZAR1bp2559 | GGTTCTGTGCAATGGTGTTTTCAT |
| Cloning ZED1 from ecotype | 9-5'-ZED1XhoIATG | GCCCTCGAGATGAGCAAGAACAATAAGAAGAAGAGA |
| | 317-3'ZED1+NOSTOP+C | CAGAGAGTTTCTCAATCAATTTGAGCT |
| NbZAR1-VIGS construct | 713-5'VIGS-NbZAR1-EcoRI | AAAGAATTCATGGTGGATGCGGTGG |
| | 714-3' VIGS-NbZAR1-1-XhoI | AAACTCGAGCTCCTCAATATGCTTCTCATG |
| q-RT-PCR | 1546-5'-TIP41 | GTGAAAACTGTTGGAGAGAAGCAA |
| | 1547-3'-TIP41 | TCAACTGGATACCCTTTCGCA |
| | 1548-5'-Fbox | GGCTGAGAGGTTCGAGTGTT |
| | 1549-3'-Fbox | GGCTGTTGCATGACTGAAGA |
| | 1620-5'ZED1-qPCR2 | TGGGTGGGGATGATGATTTGG |
| | 1621-3'ZED1-qPCR2 | TTCAAGCATACCACTCACTGGA |
| | 1552-5'-ZAR1qPCR | AGTCCATCGTGGTTGAGTCC |
| | 1553-3'-ZAR1qPCR | CCAGAATGGTTCCTGCATTT |
| Single nucleotide mutations | 1111-5'ZAR1K195N | GGCTAGGTAACACTACAATCG |
| | 1112-3'ZAR1K195N | CTCCCATCCCAACAAAGG |
| | 513-ZAR1V202M-F | CGCTCAAGAAATGTTCAATGAC |
| | 514-ZAR1V202M-R | ATTGTAGTTTTACCTAGCC |
| | 517-ZAR1S291N-F | CAAGGAGGTAACGTTATCGTG |
| | 518-ZAR1S291N- R | TCCTCTAGGCAGTCCTT |
| | 519-ZAR1P359L-F | AAAGGTTTGCTTTTGACAATCAAAG |
| | 520-ZAR1P359L-R | GCATTTCGTTACAATCTCC |
| | 521-ZAR1L465F-F | CTTTTCCGGGTTCACAAACAG |
| | 522-ZAR1L465F-R | CAGTCTTCTCCAGATTCAG |
| | 525-ZAR1G645E-F | ACGAACTGTGAATCCCTTGAATG |
| | 526-ZAR1G645E-R | CATGTCAAGTACAAGGAGTTTC |
| | 527-ZAR1P816Q-F | CAGTCAATGCTGTACCTGAGG |
| | 528-ZAR1P816Q-R | TTGCAGCACTTCCCAGTC |
| | 529-ZAR1S831F-F | GAGCTAGAGTTTTTCGCGATTGAAG |
| | 530-ZAR1S831F-R | TGGACACCAGTTTGCGGT |
| | 820-ZAR1R211C-F | GATCGAACATTGTTTCGAAAG |
| | 821-ZAR1R211C-R | TCTTTGTCATTGAACACTTC |
| | 822-ZAR1M230L-F | AGAACAGATCCTGAGAAGCATATTG |
| | 823-ZAR1M230L-R | TCAGTAAAAGTCTGGGAC |
| | 824-ZAR1L258M-F | GCAGTATCTTATGGGGAAAAG |
| | 825-ZAR1L258M-R | TGGATTTTTCTCAGCAATG |
| | 826-ZAR1S584R-F | AGATTGCAAGGCTTCAACACTTAG |
| | 827-ZAR1S584R-R | CATCCAAAATCTCAGACAGAG |
| | 1084-5'ZAR1A669G | GGGATTTAAGCCAGGCAGGTCCAACAAT |
| | 1085-3'ZAR1A669G | ATTGTTGGACCTGCCTGGCTTAAATCCC |
| | 1086-5'ZAR1N682K | CTCAGTGAAGTGAAAAAACTCACAAACCTTCGG |
| | 1087-3'ZAR1N682K | CCGAAGGTTTGTGAGTTTTTTCACTTCACTGAG |

| | | |
|---|---|---|
| | 832-ZAR1G774S-F | AATATGTTCAAGTAATTTGGTGAAAATG |
| | 833-ZAR1G774S-R | GACATGTACCTCAACATC |
| | 1088-5'ZAR1Q813P | GGGAAGTGCTGCAACCGTCAATGCCGTACC |
| | 1089-3'ZAR1Q813P | GGTACGGCATTGACGGTTGCAGCACTTCCC |
| | 1256-5'ZAR1ΔH438 | GGATGGATCGGAGAAGGT |
| | 1257-3'ZAR1ΔH438 | TACTAGTTGTTGTTTTGGTATAAC |
| Cloning ZAR1-LRR | 1191-5'ZAR1-XhoI505 | GCGCTCGAGATGTCTAACCCGGAAGGCCTCAAC |
| Cloning ZAR1 under native promoter | 18-5'ZAR1-1000ntattB | AAAGCAGGCTATAGCATTACTAATGTTTTGGTTAGTCA |
| | 180-3'ZAR1bp432 | ATCTCTTCCAACATTACTCGGCGT |
| | 56-5'ZAR1ATG | ATGGTGGACGCTGTTGTAACAGTG |
| | 19-3'MycattB | GAAAGCTGGGTGTTATCGATTTCGAACCCGGGGTAC |
| | 53-5'F attB1 complete | GG GGA CAA GTT TGT ACA AAA AAG CAG GCT |
| | 54-3'R attB2 complete+G | GGGGACCACTTTGTACAAGAAAGCTGGGTG |
| | 68-5'SalIattB1 | AGCGGTCGACATAGGGGACAAGTTTGTACAAAAAAGCAG |
| | 69-3'NotIattB2 | ATAGAAGCGGCCGCAAGGGGACCACTTTGTACAAGAAAGC |
| Cloning At5g48620 | 1508-5'XhoIAt5g48620 | GCGCTCGAGATGGCTGAAGGATTTGTGTCGTT |
| | 1509-5'XhoILRR48620 | GCGCTCGAGATGACTTCCACCTCTACCATCAATGC |
| | 1510-3'At5g48620 | CAGGTCACAGTTGATAAATTGAACATCG |

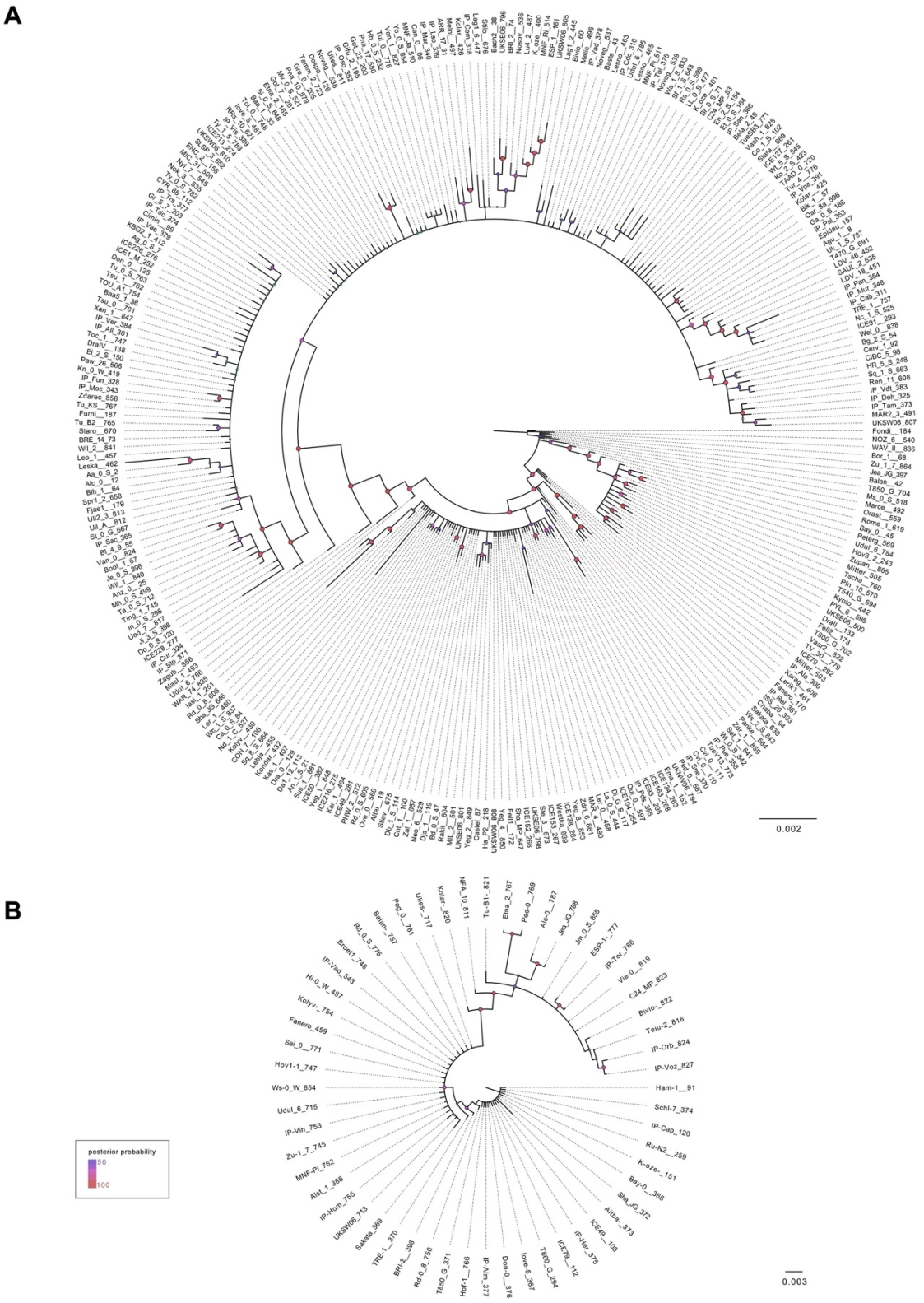



**Figure S1. Phylogenetic trees of *ZAR1* and *ZED1*.**A. MrBayes maximum clade credibility tree for ZAR1 with branches of less than 50 percent posterior probability collapsed. B. MrBayes maximum clade credibility tree for ZED1 with branches of less than 50 percent posterior probability collapsed.

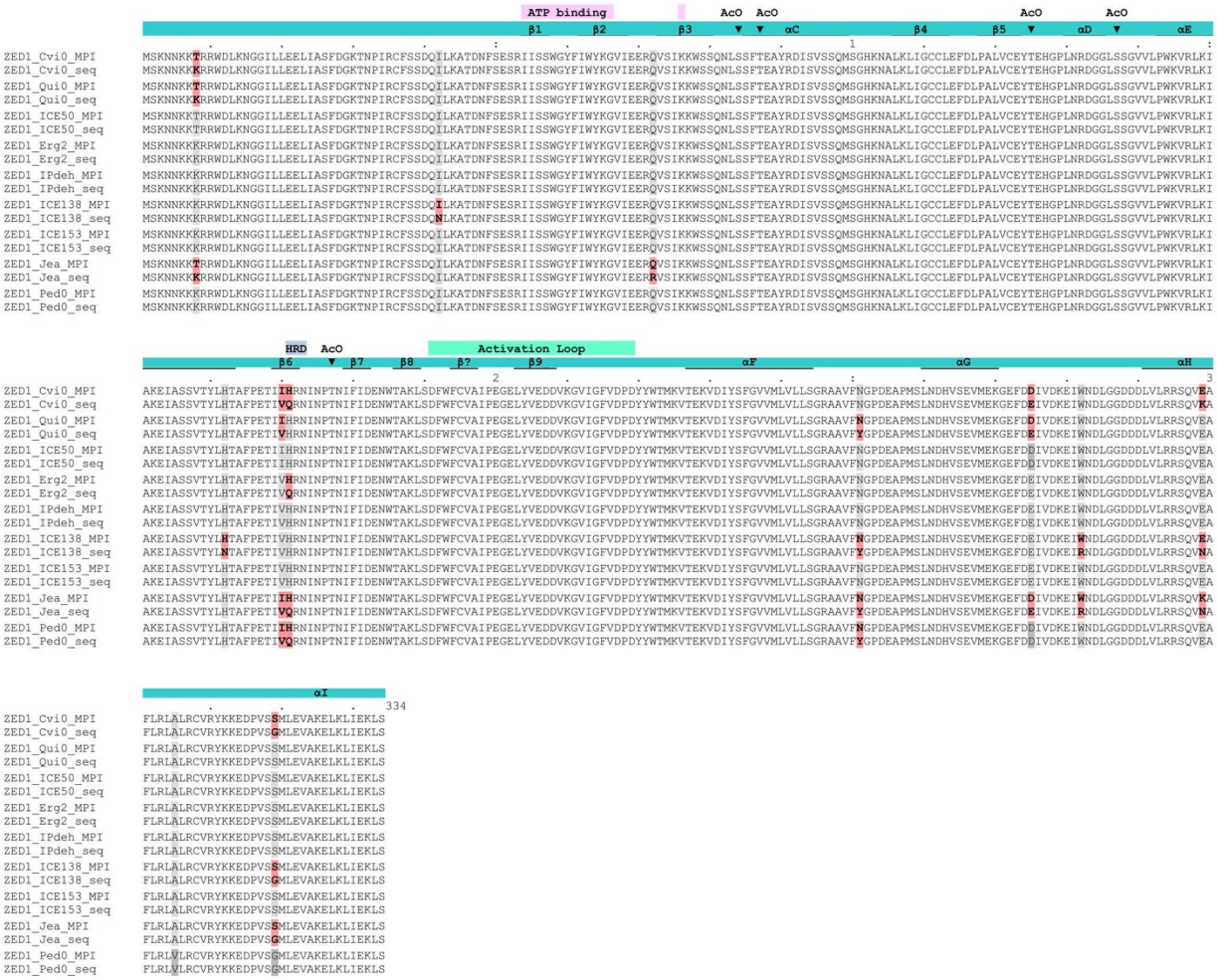



**Figure S2. Alignments of ZED1 sequences from ecotypes.** ZED1 protein sequences were extracted from the 1001 genome database (annotated with -MPI extension) and compared to the cloned sequence from each of the nine candidate ecotypes (annotated with -seq extension). The alignment displays a color-coded consensus sequence (white - conserved in all displayed sequences, light gray - moderately variable, dark gray - highly variable), while differences seen between the MPI and cloned sequences of the same ecotype are highlighted in red. Secondary structure elements [beta sheets (β) and helical segments (α), using the nomenclature from Bastedo et al., 2019], are shown above the alignment, as well as sites acetylated by HopZ1a (AcO).

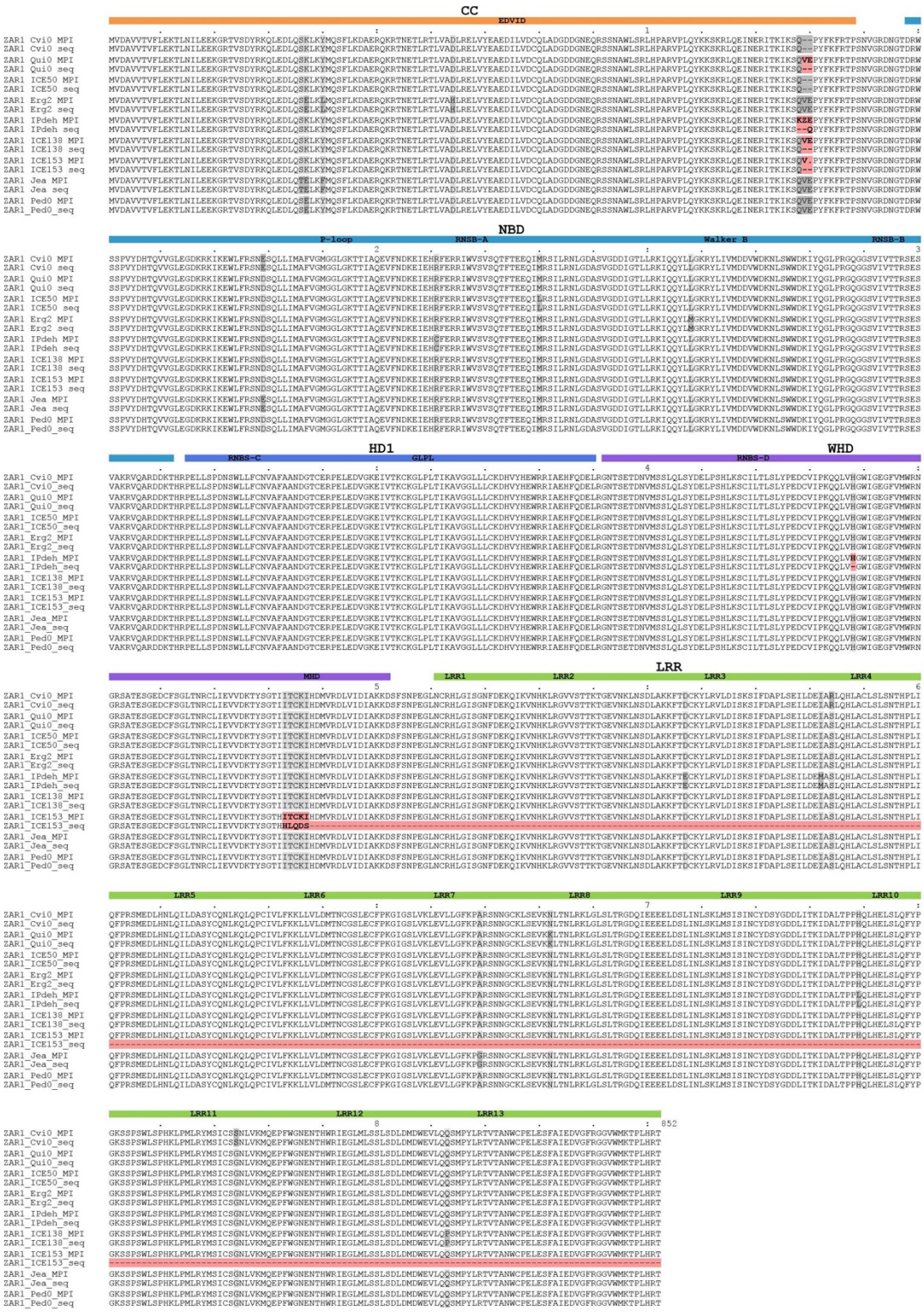



**Figure S3. Alignments of ZAR1 sequences from ecotypes.** ZAR1 protein sequences were extracted from the 1001 genome database (annotated with -MPI extension) and compared to the cloned sequence from each of the nine candidate ecotypes (annotated with -seq extension). The alignment displays a color-coded consensus sequence (white - conserved in all displayed sequences, light gray - moderately variable, dark gray - highly variable), while differences seen between the MPI and cloned sequences of the same ecotype are highlighted in red. Domains are colored as follows: CC - orange; medium blue - NBD , dark blue - HD1, purple - WHD, and LRR - green.

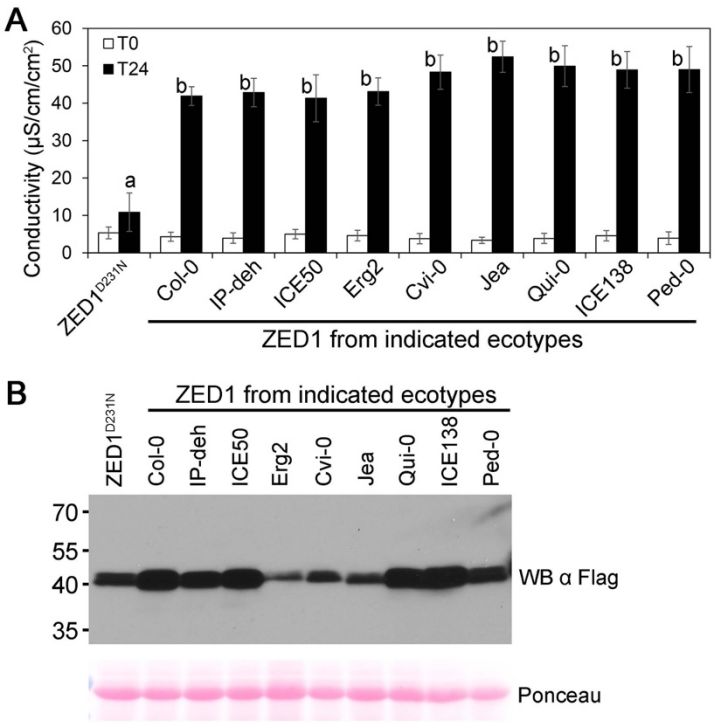



**Figure S4. ZED1 alleles from ecotypes are not impaired in HopZ1a recognition.** A. Levels of ion leakage induced by the co-expression of HopZ1a and ZED1$^{D231N}$ or ZED1 alleles from ecotypes, were measured at 4 h (T0) and 28 h (T24) after dexamethasone application. Error bars indicate the standard error from the mean of six samples. The letters beside the bars indicate significance groups, as determined by a one-way analysis of variance (ANOVA) comparison followed by a Tukey's post hoc test ($P \leq 0.05$). The experiment was repeated twice with similar results. B. *N. benthamiana* tissues expressing *ZED1* alleles were harvested 24 h after dexamethasone induction. ZED1 protein expression was detected with an α-Flag antibody. Ponceau staining of total protein extract was used as loading control.

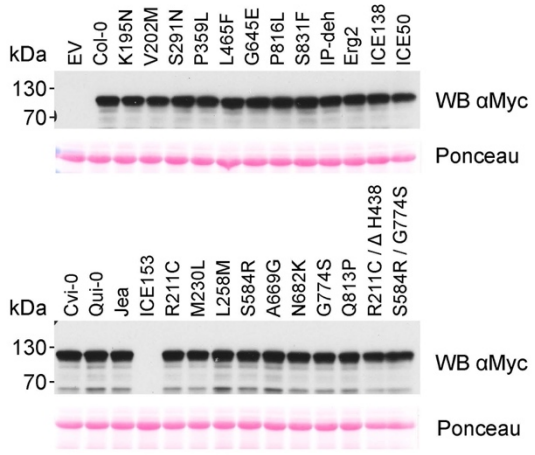



**Figure S5. Protein expression of ZAR1 constructs in *N. benthamiana*.** *N. benthamiana* tissues expressing *ZAR1* constructs were harvested 24 h after dexamethasone induction. ZAR1 protein expression was detected with an α-Myc antibody. Ponceau staining of total protein extract was used as a loading control.

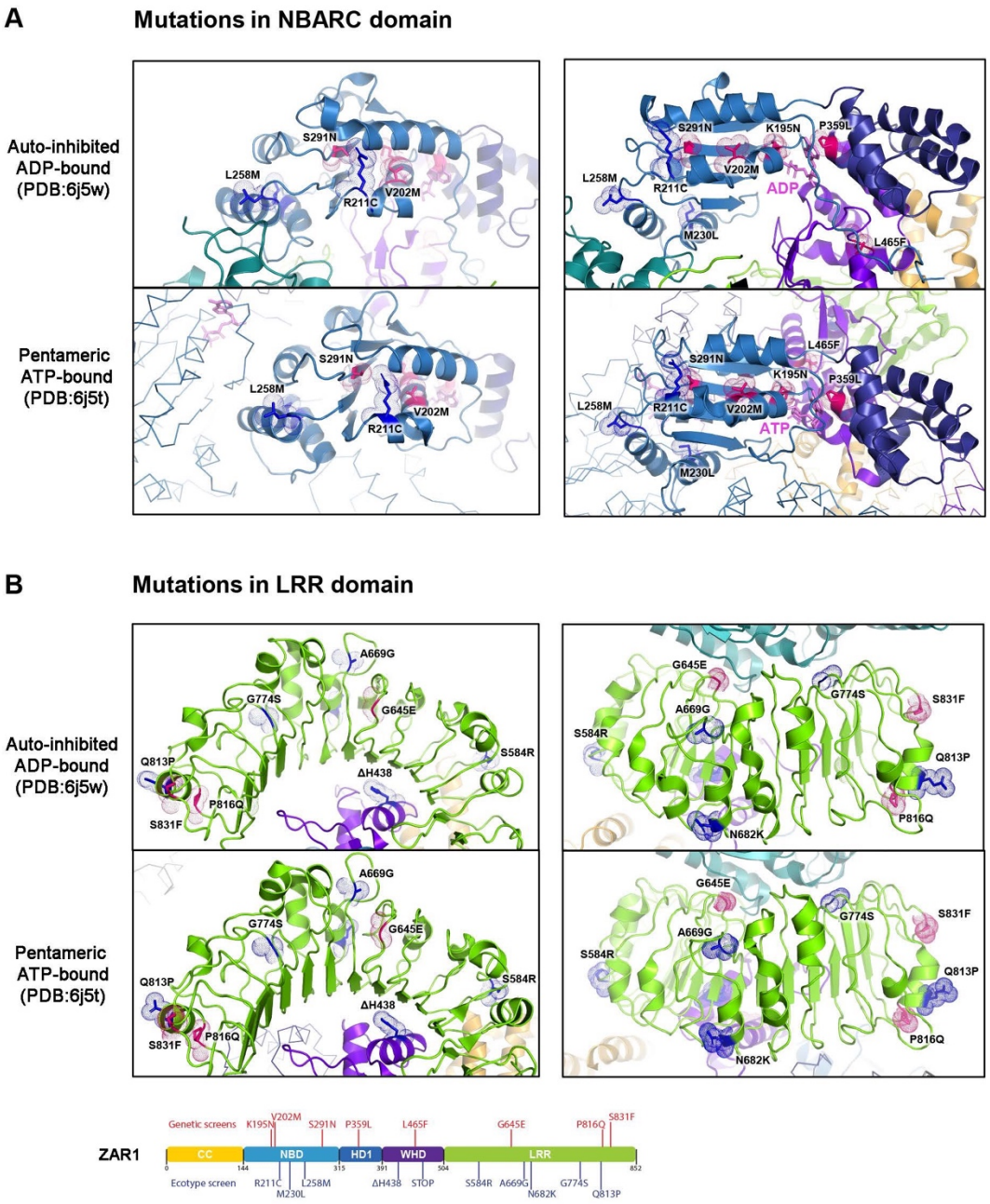



**Figure S6. Polymorphisms in ZAR1 mapped onto cryo-EM structures.** ZAR1 NBARC (A) and LRR (B) mutations mapped onto the cryo-EM structures of both the auto-inhibited ADP-bound conformation (PDB: 6j5w) and the activated ATP-bound pentamer (PDB: 6j5t). Ecotype polymorphisms (blue) and genetic screen (red) mutations are illustrated with sticks & dots representation. ZAR1 domains are colored as indicated at the bottom of the figure: CC – orange; NBD – blue, HD1 – dark blue, WHD – purple, LRR – green, RKS1 kinase – cyan and ADP/ATP – pink. Neighboring ZAR1 monomers from the pentameric conformation are shown using ribbon representation (thin lines). Graphics were generated using The PyMOL Molecular Graphics System, Version 2.2.3 Schrödinger, LLC.

## CONCLUSION

By identifying key residues responsible for ZAR1 recognition of the *P. syringae* effector HopZ1a, this work is a key step towards rational design of ZAR1 recognition pathways for *P. syringae* resistance.

# CHAPTER III: Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1

## PREFACE

This chapter was published in 2023 as "Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1" in *Plant, Cell, & Environment*. I performed the transient gene expressions, co-immunoprecipitations, selection analyses, binding interface analysis, analyzed the data, and wrote the manuscript with assistance from all authors. The publication has been reproduced below with the permission of all authors.

This work uses natural diversity in ZAR1 interacting proteins to identify determinants for ZAR1 interaction specificity. This is a key step towards rational design of ZAR1 recognition pathways for *P. syringae* resistance.

Diplock N., Baudin M., Harden L., Silva C.J., Erickson-Beltran M.L., Hassan J.A., Lewis J.D. Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1. *Plant Cell Environ* (46) 2238–2254 (2023).

**Abstract**

The highly conserved angiosperm immune receptor HOPZ-ACTIVATED RESISTANCE1 (ZAR1) recognizes the activity of diverse pathogen effector proteins by monitoring the ZED1-related kinase (ZRK) family. Understanding how ZAR1 achieves interaction specificity for ZRKs may allow for expansion of the ZAR1-kinase recognition repertoire to achieve novel pathogen recognition. We took advantage of the natural diversity of *Arabidopsis thaliana* kinases to probe the ZAR1-kinase interaction interface and found that ZAR1 can interact with most ZRKs, except ZRK7. We found evidence of alternative splicing of ZRK7, resulting in a protein that can interact with ZAR1. Despite high sequence conservation of ZAR1, interspecific ZAR1-ZRK pairings resulted in auto-activation of cell death. We showed that ZAR1 interacts with a greater diversity of kinases than previously thought, while still possessing the capacity for specificity in kinase interactions. Finally, using ZAR1-ZRK interaction data, we rationally increased ZRK10 interaction strength with ZAR1, demonstrating the feasibility of rational design of a ZAR1-interacting kinase. Overall, our findings advance our understanding of the rules governing the ZAR1 pathogen recognition system, with promising future directions for expanding ZAR1 immunodiversity.

**Introduction**

Plant pathogens present a significant threat to global agriculture and food security. Emerging pathogens that evade population level immunity may require novel approaches to achieve resistance. Plants achieve immunity against pathogens through two interconnected layers of defense [1–3]. Initial resistance occurs at the cell surface, where pattern recognition receptors bind pathogen molecules and initiate pattern-triggered immunity (PTI). Some bacterial pathogens including *Pseudomonas syringae* and *Xanthomonas* spp. can secrete proteins known as type III secreted effector proteins (T3SEs) into the plant cell, where they promote virulence and suppress PTI. Plants have intracellular nucleotide-binding leucine-rich repeat receptors (NLRs) that can recognize pathogen effectors, resulting in effector-triggered immunity (ETI). NLRs recognize effectors either through direct NLR-effector interaction, or by "guarding" for

effector modification of host proteins [4]. A hallmark of ETI is a macroscopic programmed cell death known as the hypersensitive response (HR).

HOPZ-ACTIVATED RESISTANCE1 (ZAR1) is a highly conserved NLR capable of recognizing diverse T3SEs from several bacterial species [5]. The molecular mechanisms of ZAR1-mediated ETI have been explored extensively in *Arabidopsis thaliana*. ZAR1 recognizes several T3SEs through modification of ZAR1-associated receptor-like cytoplasmic kinases (RLCKs). Several members of the RLCK-XII family of pseudokinases, including HOPZ-ETI-DEFICIENT1 (ZED1), and ZED1-related kinases (ZRKs) interact with ZAR1 in the absence of pathogen effectors [6–9]. Some T3SEs modify kinases in the RLCK-VII family, including PBS1-like kinases (PBLs) to trigger PBL-ZRK interactions and ultimately ZAR1 activation [7,10]. However, other T3SEs such as HopZ1a modify ZRKs directly leading to activation independent of PBLs [6]. Overall, ZAR1 achieves recognition of diverse T3SEs by guarding an initial layer of ZRKs which interact with bacterial T3SEs either directly or through a second layer of PBLs.

Several NLRs "guard" host proteins to detect modification by T3SEs [11–16]. For example, the *A. thaliana* NLR RPS5 guards AVRPPHB SUSCEPTIBLE 1 (PBS1), and triggers ETI when PBS1 is cleaved by *P. syringae* effector AvrPphB [16]. However, unlike many NLRs, ZAR1 can respond to a unique breadth of T3SEs that have different enzymatic functions and that come from several bacterial species. ZAR1 immunodiversity is facilitated through the guarding of multiple ZRKs [6,9]. In *A. thaliana,* the ZRK family consists of 13 members with ZED1, RKS1, ZRK2 and ZRK3 having a demonstrated role in effector recognition [6–9]. The *Pseudomonas syringae* T3SE HopZ1a directly interacts with and acetylates ZED1, triggering ETI though ZAR1 [6]. The *P. syringae* T3SE HopX1i is also recognized through ZED1, but requires an additional RLCK-VII, Suppressor of ZED1-D (SZE1) for ZAR1-mediated ETI [9]. ZAR1 also mediates ETI in response to *P. syringae* T3SEs HopZ5 through ZED1, HopF1r, HopO1c, HopO1b, and HopO2a through ZRK3, and HopBA1 through ZRK2 [9,17,18]. While PBL27 has been shown to be required for HopF1r recognition, no PBLs have been identified for recognition of other *P. syringae* T3SEs [10]. In addition to diverse *P. syringae* effectors, ZAR1 can recognize the *Xanthomonas campestris* T3SE AvrAC, which uridylates PBL2, triggering an interaction between PBL2 and RKS1 and the activation of ZAR1 [7]. While ZED1, RKS1, and ZRK3 have been shown to interact directly with ZAR1, no direct interaction between ZRK2 and ZAR1 has been observed [6–8]. Although they have no known role in effector recognition, ZRK6 and ZRK15 have also been shown to interact with ZAR1 [7]. A recent detailed review of the ZAR1 effector recognition system highlights these effector recognition pathways in additional detail [19].

In addition to mediating extensive immunodiversity within *A. thaliana*, ZAR1 is functionally conserved across angiosperms. We have previously shown that the *Nicotiana benthamiana* ZAR1 ortholog (hereafter NbZAR1) can recognize HopZ1a when *A. thaliana* ZED1 is transiently expressed in *N. benthamiana* [20]. In addition, a ZRK homolog in *N. benthamiana* mediates *X. perforans* effector recognition through NbZAR1 [21]. Furthermore, transiently overexpressing AtRKS1 with ZAR1 orthologs from multiple eudicot species results in autoactive cell death of *A. thaliana* protoplasts [22]. This suggests that ZAR1 guarding of ZRKs is conserved across angiosperms. Interspecific pairings of ZAR1 and ZRKs may constitute a functional pathogen recognition system or lead to autoactivation of ETI and programmed cell death.

ZAR1 is a canonical NLR comprised of an N-terminal coiled-coil domain (CC, hereafter $ZAR1^{CC}$), a central nucleotide binding domain (NBD, hereafter $ZAR1^{NBD}$) and a C-terminal leucine-rich repeat domain (LRR, hereafter $ZAR1^{LRR}$; Baudin et al., 2019; Lewis et al., 2010).

The cryo-EM structure of ZAR1 during three stages of activation upon AvrAC recognition was a breakthrough for our understanding of ZAR1-ZRK interactions and ZAR1 molecular function. In the absence of *X. campestris* effector AvrAC, RKS1 interacts with ZAR1 primarily through $ZAR1^{LRR}$ [24,25]. We have previously demonstrated that $ZAR1^{LRR}$ is required and sufficient for ZED1 interaction *in planta* [20], which supports the importance of $ZAR1^{LRR}$ for ZRK interaction. When PBL2 is modified by AvrAC, a PBL2-RKS1-ZAR1 complex forms, resulting in stabilization of the RKS1 activation loop which then allosterically clashes with $ZAR1^{NBD}$ [24]. This clash triggers extensive conformational changes, causing ZAR1 to oligomerize with other activated ZAR1 complexes to form a wheel-like pentamer, termed the "resistosome" [25]. In the resistosome conformation, RKS1 interacts with ZAR1 exclusively through $ZAR1^{LRR}$ [25]. While the ZAR1 structure is invaluable for understanding the ZAR1-RKS1 interaction, the determinants of ZAR1 specificity for the broader ZRK family are not well understood.

Plant kinase repertoires are generally larger than those in other Eukaryotes and are common targets of T3SEs [26–29]. The *A. thaliana* proteome is made up of almost 5% kinases [30], and yet ZAR1 appears to achieve specificity for a small subset of the Arabidopsis kinase repertoire [31]. Understanding the determinants of ZAR1 specificity for kinase guardees may allow for the identification of additional ZAR1-mediated pathogen recognition components, as well as the rational design of a ZAR1 guardee as a first step towards generating a novel ZAR1 effector recognition pathway. Here, we show that ZAR1 can achieve interaction specificity for kinases both within and outside of the *A. thaliana* ZRK family (hereafter ZRKs). We also show that interspecific pairing of ZRKs with NbZAR1 results in an autoactive HR-like phenotype. We demonstrate the feasibility of rational guardee design by increasing the interaction strength between $ZAR1^{LRR}$ and ZRK10. Overall, we identify key determinants of ZAR1-kinase interaction specificity using natural diversity of kinases.

**Results**

*ZAR1 interacts with multiple ZRKs under native conditions in A. thaliana*

To determine which ZRKs interact with ZAR1 during immune activation, we transformed *A. thaliana zar1-1* plants with a native promoter-driven *ZAR1* construct carrying an in-frame 3xFlag epitope tag (hereafter *pZAR1:ZAR1-3xFlag*). We also generated control lines with a native promoter driven *ZAR1* construct lacking the 3xFlag tag (hereafter *pZAR1:ZAR1*). We screened T1 and T2 plants for Basta resistance (Supplemental Figure S1), presence of the transgene, the *zar1-1* genotype, protein expression in the tagged lines and complementation of HopZ1a recognition when plants were pressure-infiltrated with *P. syringae* pv. *tomato* DC3000 (*Pto*DC3000) carrying *hopZ1a*. We confirmed that ZAR1 protein expression was maintained in the T3 lines and showed that the lines were still complemented for the HopZ1a-triggered HR (Figure 1A).

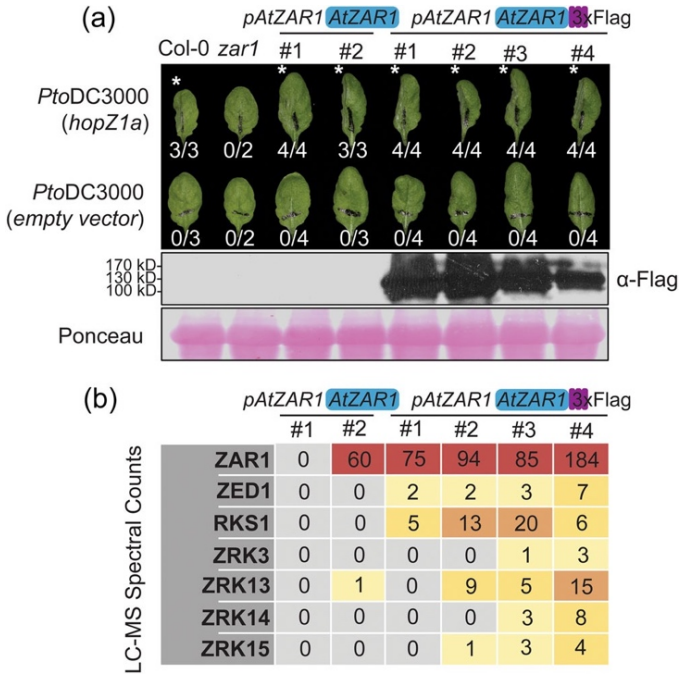



**Figure 1.** LC-MS identifies ZRKs that interact with ZAR1 under native conditions in *A. thaliana*. (A) *A. thaliana zar1* plants expressing native promoter *ZAR1* with a 3xFlag epitope tag (*pZAR1:ZAR1-3xFlag*) or untagged native promoter *ZAR1* (*pZAR1:ZAR1*) show complementation of HopZ1a recognition. Half-leaves of T3 lines were infiltrated with *Pto*DC3000 carrying *hopZ1a* or an empty vector. Asterisks next to leaves indicates the presence of the hypersensitive response. Numbers under leaves indicate the number of leaves with a strong HR out of total infiltrated leaves. Anti-flag western blots show that ZAR1-3xFlag is detected in *pZAR1:ZAR1-3xFlag* plants, but not *pZAR1:ZAR1* plants. (B) Liquid chromatography tandem mass spectroscopy (LC-MS/MS) show an enrichment of RKS1, ZRK3, ZRK13, ZRK14, and ZRK15 spectral counts in *pZAR1:ZAR1-3xFlag* plants relative to *pZAR1:ZAR1* lines. Tissue samples for LC-MS/MS were taken 3 days after plants were spray inoculated with *Pto*DC3000 carrying *hopZ1a*. Red indicates spectral counts greater than 20, orange greater than 10, yellow greater than 0, and grey equal to 0.

Next, we selected four *pZAR1:ZAR1-3xFlag* lines and two untagged *pZAR1:ZAR1* control lines and collected tissue samples from 6–7-week-old plants, 3 days after spray inoculation with *Pto*DC3000 (*hopZ1a*). We conducted immunoprecipitation followed by liquid chromatography mass spectrometry (LC-MS). We searched the spectral counts for kinases, and found an enrichment of ZED1, RKS1, ZRK3, ZRK13, ZRK14 and ZRK15 in *pZAR1:ZAR1-3xFlag* lines compared to *pZAR1:ZAR1* lines (Figure 1B). Our results indicate that ZRK13 and ZRK14 can interact with ZAR1, suggesting that ZAR1 can interact with additional ZRKs beyond those already reported [6–8]. To further extend these findings, we decided to investigate ZAR1-ZRK interactions through transient expression in *N. benthamiana*.

*ZAR1 displays interaction specificity within the ZRK family*

To further investigate whether ZAR1 displays interaction specificity for kinases within the ZRK family, we cloned all *ZRK*s (Figure 2A) with 3x-Flag fusions and transiently overexpressed each ZRK with either ZAR1 or ZAR1$^{LRR}$ fused to a 5x-Myc epitope tag in *N. benthamiana*. We collected tissue samples and conducted co-immunoprecipitation (co-IP) experiments to test for ZAR1-ZRK and ZAR1$^{LRR}$-ZRK interactions. As a negative control for ZAR1 and ZAR1$^{LRR}$ interaction we cloned *BOTRYTIS-INDUCED KINASE1* (*BIK1*), a kinase from the PBL family which has previously been shown to not interact with ZAR1 [7], with a 3x-

Flag fusion. We observed an interaction between ZAR1 and all ZRKs, except for ZRK7 (Figure 2B), and between ZAR1$^{LRR}$ and all ZRKs, except for ZRK7 (Figure 2C). Our findings indicate that ZAR1 and ZAR1$^{LRR}$ have identical interaction specificities and that ZAR1$^{LRR}$ is sufficient for ZRK interaction.

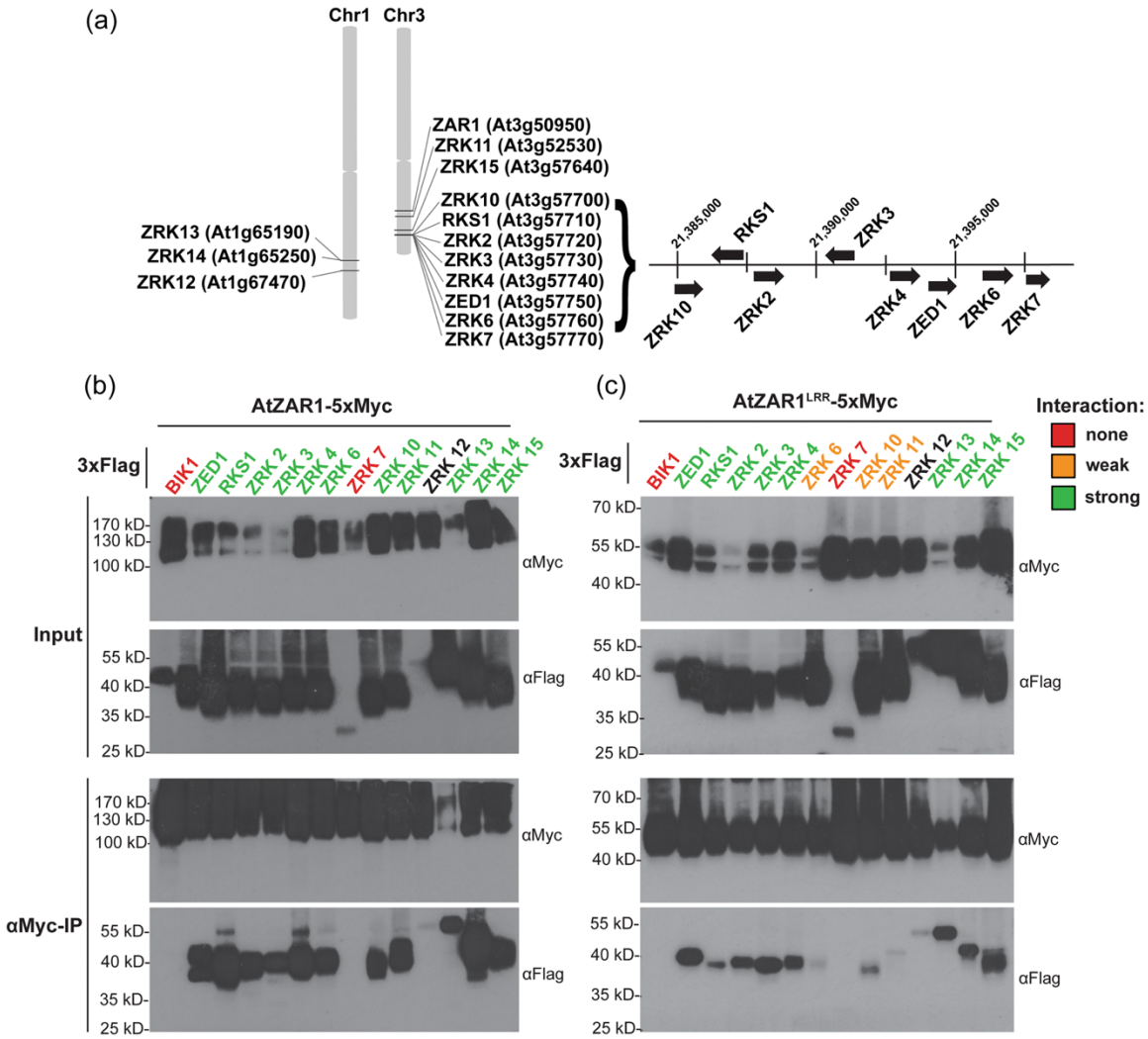



**Figure 2.** ZAR1 and ZAR1$^{LRR}$ achieve interaction specificity within the ZRK family. (A) Depiction of *ZRK* locations in the *A. thaliana* genome, with the genes in the main *ZRK* cluster shown in detail. Representative western blots of ZAR1 (B) and ZAR1$^{LRR}$ (C) co-immunopurification with ZRKs in *N. benthamiana*. Each ZRK was fused with a 3xFlag tag and co-expressed with ZAR1 or ZAR1$^{LRR}$ which were fused with a 5xMyc tag. Anti-Myc beads were used to immunoprecipitate ZAR1 and ZAR1$^{LRR}$. Green text indicates a strong interaction, yellow indicates a weak interaction and red indicates a lack of interaction. ZRK12 is colored black since the protein was consistently harder to detect compared to other ZRKs and we did not make a conclusion on the strength of the ZRK12 interactions as a result. All experiments were repeated three times with similar results. Predicted molecular masses of the proteins are ZAR1-5xMyc 107 kD, ZAR1$^{LRR}$-5xMyc 50kD; with 3xFlag: BIK1 47kD, ZED1 41kD, RKS1 43kD, ZRK2 44kD, ZRK3 44kD, ZRK4 44kD, ZRK6 46kD, ZRK7 33kD, ZRK10 42kD, ZRK11 42kD, ZRK12 48kD, ZRK13 49kD, ZRK14 46kD, and ZRK15 44kD.

We consistently observed the strongest interactions between ZAR1$^{LRR}$ and ZED1, RKS1, ZRK3, ZRK4, ZRK13, and ZRK14. Notably, these strong ZAR1$^{LRR}$ interactors include all ZRKs with a previously characterized role in effector recognition through ZAR1, supporting that a

strong ZAR1$^{LRR}$ interaction is biologically relevant for effector recognition [6–9,20,32]. Moreover, the ZRKs that interact strongly with ZAR1$^{LRR}$ overlap with the subset of ZRKs identified by LC-MS in *A. thaliana* (Figure 1B). We observed a consistently weak interaction between ZAR1$^{LRR}$ and ZRK6, ZRK10, and ZRK11 (Figure 2B). While the interaction between ZAR1 and ZRK12 was noticeably weaker when compared to other ZRKs, the ZRK12 protein was consistently harder to detect compared to other ZRKs (Figure 2B, C). Thus, we cannot make a conclusion on the strength of the ZAR1-ZRK12 and ZAR1$^{LRR}$-ZRK12 interaction relative to the other ZRKs.

*Arabidopsis ZRKs cause NbZAR1-dependent cell death in N. benthamiana*

While transiently overexpressing ZRKs in *N. benthamiana* for co-IP experiments, we noticed that some ZRKs resulted in an HR in the absence of pathogen effectors. To determine if the HR is *NbZAR1*-dependent, we used virus-induced gene silencing (VIGS) to target *NbZAR1* (hereafter *NbZAR1*-VIGS) or *GUS* (hereafter *GUS*-VIGS) as a negative control [20]. We overexpressed each ZRK in *NbZAR1*-VIGS and *GUS*-VIGS plants. We scored the hypersensitive response 24 hours post-dexamethasone induction as either no HR, weak HR, medium HR, or strong HR (Figure 3A). Both RKS1 and ZRK13 resulted in a strong *NbZAR1*-dependent autoactive HR phenotype, while the other ZRKs did not (Figure 3B). To ensure *NbZAR1* was silenced, we overexpressed HopZ1a and ZED1 together on each leaf used across all experiments. We did not observe an HR in *NbZAR1*-silenced plants, indicating that *NbZAR1* silencing was effective. As previously observed [20], we consistently observed an HR when AtZAR1 was co-expressed with ZED1 and HopZ1a, since it could complement *NbZAR1* silencing, and did not see an HR when ZED1 was expressed alone, since ZED1 alone is not autoactive (Figure 3B). Therefore, of the 13 tested ZRKs, only RKS1 and ZRK13 caused a strong *NbZAR1*-dependent autoactive HR phenotype.

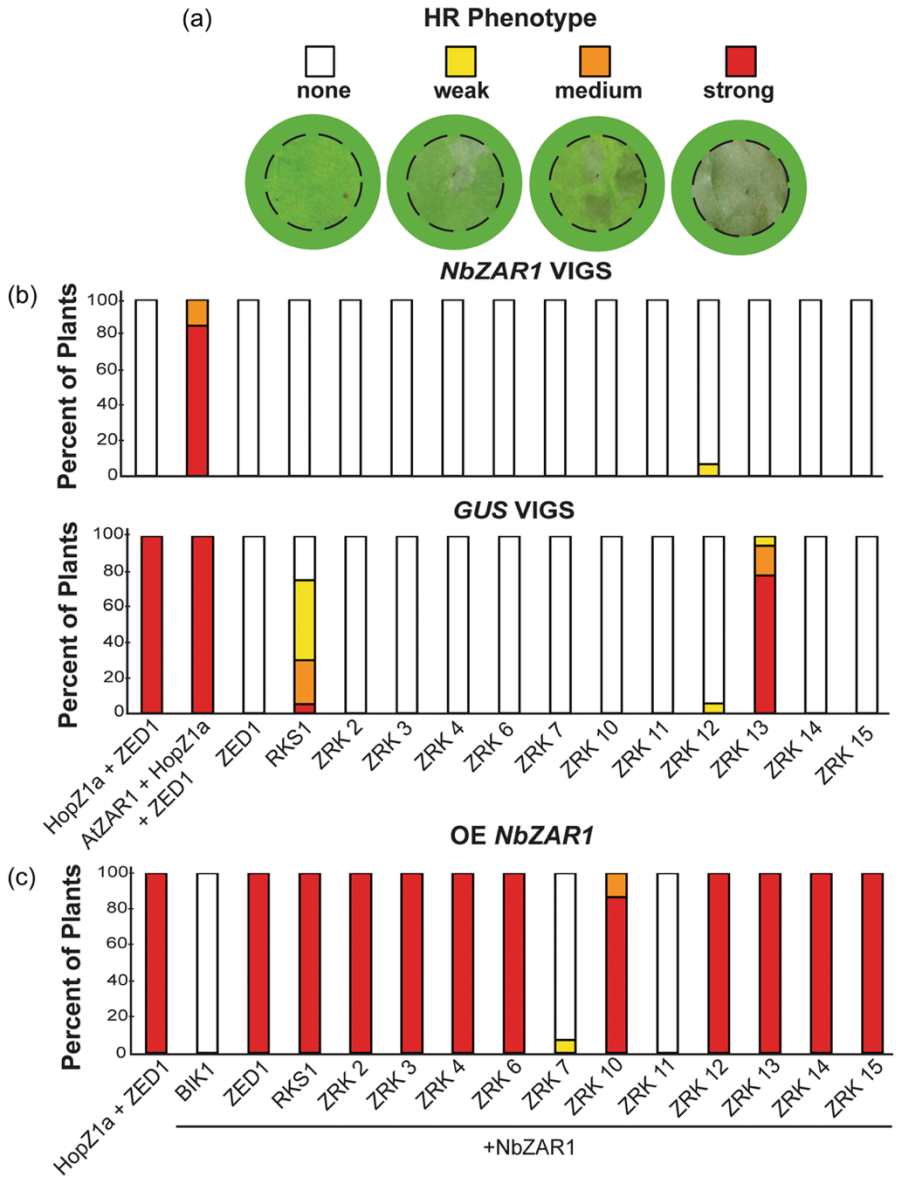



**Figure 3.** Several *A. thaliana* ZRKs result in *N. benthamiana* ZAR1-dependent HR in the absence of pathogen effectors. (A) HRs were scored on a 4-category scale depicted by representative images on a leaf-colored background. (B) All ZRKs were transiently overexpressed in *N. benthamiana* plants silenced for either *NbZAR1* or *GUS* and infiltrated spots were scored for the HR 24 hours after dexamethasone induction (*NbZAR1* n=14-17, *GUS* n=18-20). Virus-induced gene silencing (VIGS) was used to target *NbZAR1* or *GUS* (Baudin et al., 2017). HopZ1a and ZED1 were used to confirm silencing of *NbZAR1*. AtZAR1, HopZ1a and ZED1 were used as a positive control for HR, and ZED1 alone was used as a negative control for HR (Baudin et al., 2017). (C) ZRKs and NbZAR1 were transiently overexpressed in *N. benthamiana* plants and scored for HR based on a 4-category scale (n=12-15). BIK1 was overexpressed with NbZAR1 as a negative control for the HR. Data is compiled from at least 3 separate experiments.

To determine if the autoactive HR is due to interspecific pairings of ZAR1 and ZRKs, we transiently overexpressed AtZAR1 and each ZRK in *NbZAR1*-VIGS plants. We did not see any strong or consistent HRs when AtZAR1 was overexpressed with each ZRK, suggesting that the observed autoactive HR is a result of interspecific ZAR1-ZRK pairings (Supplemental Figure S2). However, we did notice a weak inconsistent HR phenotype when ZRK3 and ZRK13 were

co-expressed with AtZAR1 (Supplemental Figure S2). *NbZAR1* was silenced effectively since co-expression of HopZ1a and ZED1 did not result in an HR. As previously observed (Baudin et al. 2017), the HR was complemented when we expressed AtZAR1, HopZ1a and ZED1 (Supplemental Figure S2).

We hypothesized that weaker *NbZAR1*-dependent HR phenotypes might be missed due to low levels of natively expressed *NbZAR1*. To test this, we transiently overexpressed NbZAR1 with each ZRK in *N. benthamiana* plants. Since overexpression of NLRs can often result in HR phenotypes, we co-expressed BIK1 with NbZAR1 as a negative control for potential NbZAR1 autoactivity. Each ZRK resulted in a strong HR when overexpressed with NbZAR1, except for ZRK7 and ZRK11 (Figure 3C). It is likely that the lack of the HR when ZRK7 and ZRK11 are respectively co-expressed with NbZAR1 is due to lack of interaction between these ZRKs and NbZAR1. In support of this hypothesis, ZRK7 cannot interact with AtZAR1 or AtZAR1$^{LRR}$ while ZRK11 weakly interacts with AtZAR1$^{LRR}$ (Figure 2C).

*Alternative variant of ZRK7 can interact with ZAR1$^{LRR}$ and full-length ZAR1*

ZRK7 was the only member of the ZRK family that did not interact with ZAR1 or ZAR1$^{LRR}$ (Figure 2). When we compared the ZRK7 amino acid sequence to that of the other ZRKs, we noticed that ZRK7 seems to be truncated at the C-terminus following the activation loop (Supplemental Figure S3). Furthermore, *ZRK7* is unique among the ZRK family due to the presence of an intron in the coding sequence. In the literature and in publicly available RNA-seq data [33,34], we found evidence of alternative splicing, resulting in intron retention of *ZRK7* (Figure 4A, Marquez et al., 2012). Intron retention results in readthrough into the *ZRK7* 3'UTR, which produces a protein that is similar in size and sequence to the other Arabidopsis ZRKs (Figure 5A, Supplemental Figure S3). We cloned the genomic sequence for *ZRK7* (hereafter *gZRK7*) as an in-frame fusion with a 3x-Flag epitope tag and tested it for interaction with ZAR1-5xMyc and ZAR1$^{LRR}$ via transient expression in *N. benthamiana* followed by co-IP. Surprisingly, gZRK7 can interact with both ZAR1$^{LRR}$ and full-length ZAR1 (Figure 4B). As previously observed (Figure 2), BIK1 and ZRK7 did not interact with either ZAR1$^{LRR}$ or full-length ZAR1. To determine if gZRK7 can activate the NbZAR1-mediated HR, we overexpressed NbZAR1 and gZRK7 in *N. benthamiana* and scored cell death 24 hours after gene induction based on a 4 category HR scale. We did not observe an HR when NbZAR1 was overexpressed with gZRK7 (Figure 4C).

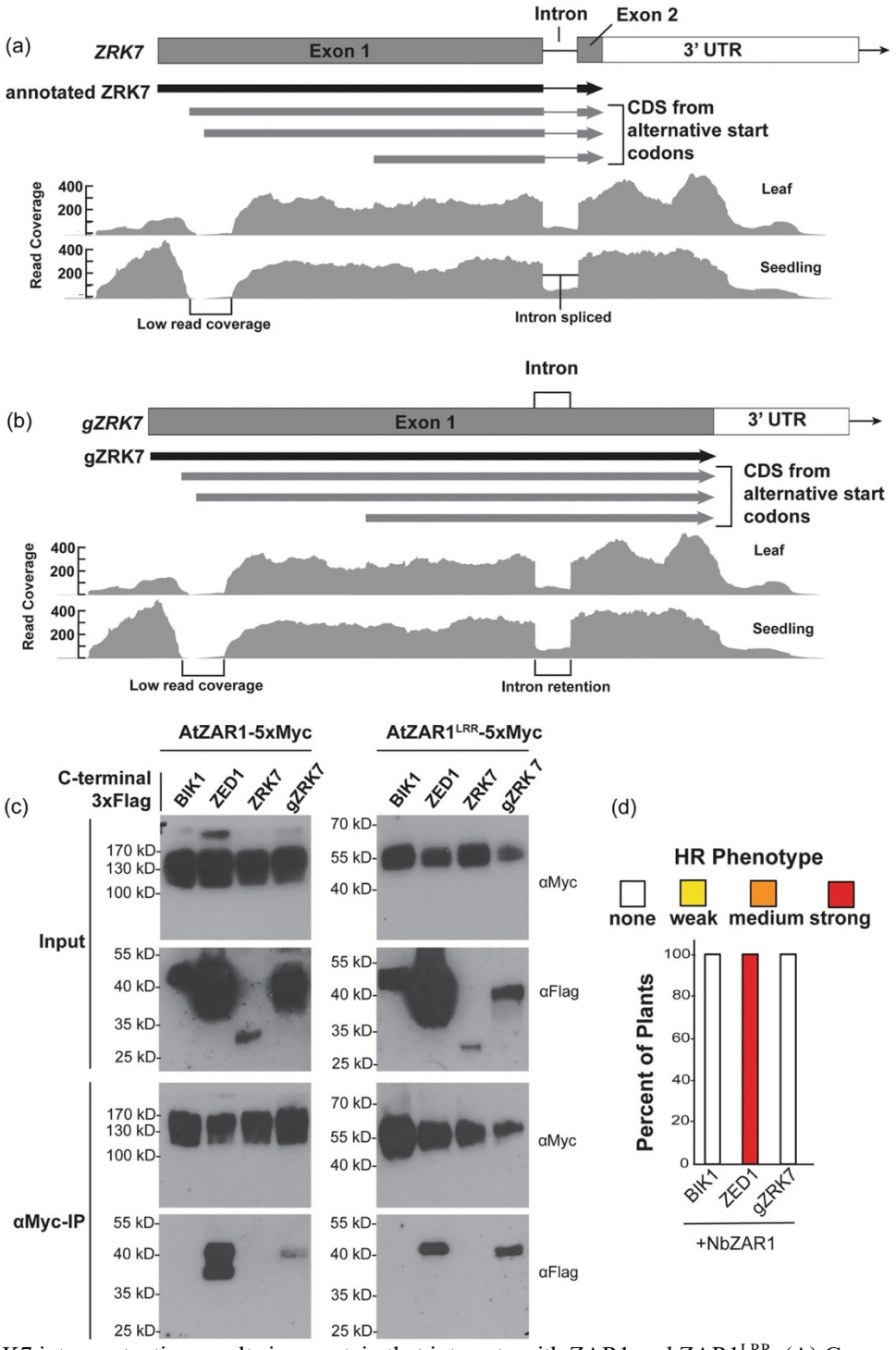



**Figure 4.** ZRK7 intron retention results in a protein that interacts with ZAR1 and ZAR1$^{LRR}$. (A) Gene annotation diagram of *ZRK7* and the intron retained genomic *ZRK7* (*gZRK7*) with grey bars representing the coding sequence, line representing the intron, black arrows representing the start codon, and grey arrows representing alternative in-frame start codons. RNA-seq read depth from publicly available data of *A. thaliana* seedling and leaf tissue are shown below gene diagrams and are aligned with the gene models (Lamesch et al., 2012; Berardini et al., 2015). (B) Representative western blots showing the co-immunoprecipitation of ZAR1 and ZAR1$^{LRR}$ with ZRK7 and gZRK7 in *N. benthamiana*. BIK1, ZED1, ZRK7 and gZRK7 were fused with a 3xFlag tag and transiently co-expressed with either ZAR1 or ZAR1$^{LRR}$ which was fused with a 5xMyc tag. Anti-Myc beads were used to immunoprecipitate ZAR1 and ZAR1$^{LRR}$. Co-immunoprecipitation experiments were repeated three times with similar results. Predicted

molecular masses of the gZRK7-3xFlag is 44kD. (C) gZRK7 and NbZAR1 were transiently overexpressed in *N. benthamiana* plants and scored for HR based on a 4-category scale (as described in Figure 3A, n=15). BIK1 was overexpressed with NbZAR1 as a negative control for the HR, while ZED1 was overexpressed with NbZAR1 as a positive control. Data is compiled from 3 separate experiments.

In publicly available RNA-seq data, we noticed relatively few reads immediately following the *ZRK7* start codon, in a pattern suggestive of an unannotated intron (Figure 4A). However, despite high read counts on either side of the region with low reads, no split reads spanning this region were identified that would support the presence of an intron in the data [33,34]. It is possible that *ZRK7* is misannotated and the transcript begins after the region of low read depth. However, the first in-frame start codon present in this transcript occurs 384 nucleotides into the annotated coding sequence. We cannot rule out the possibility that *ZRK7* is a pseudogene, or that smaller reading frames within the *ZRK7* gene other than the two tested in this study are of importance of the ZAR1 recognition system. Regardless, we have shown that *A. thaliana* possesses an unannotated *ZRK7* transcript encoding a protein that can interact with both ZAR1 and ZAR1$^{LRR}$.

We hypothesized that *ZRK7* may represent an intermediate stage in ZRK evolution and conducted a selection analysis to determine if *gZRK7* was under different selective pressure relative to other ZRKs. We tested for episodic diversifying selection across ZRKs using the adaptive Branch Site REL (aBSREL) model available through Hyphy. We saw evidence of gene level diversifying selection in *gZRK7*, as well as in *ZRK12*, *ZRK13* and *ZRK15* (Supplemental Figure S4). To identify residues under different selection pressure in *gZRK7* compared to other ZRKs, we used the contrast-FEL model. We identified 9 *gZRK7* residues with significantly different dN/dS ratios relative to other ZRKs. Interestingly, 7 of the 9 residues are located prior to the region of low RNA-seq read coverage at the *ZRK7* 5' end, suggesting that this region may be under strong selective pressure relative to the other ZRKs (Supplemental Figure S3; Supplemental Table S1).

*ZAR1 interacts with kinases outside of the ZRK family*

Our LC-MS data suggested that ZAR1-kinase interaction specificity is confined to the ZRK family. Based on trends from the LC-MS data (Figure 1B) and previously published data (Gong et al., 2022; G. Wang et al., 2015), we hypothesized that ZAR1 would interact with ZRKs, but not closely related wall associated kinase-like proteins (WAKLs) or PBLs (Figure 5A). We cloned PBL2, and the kinase domain of WAKL2 (aa 346-748, hereafter WALK2$^{KIN}$) as a fusion with a 3x-Flag tag. WALK2$^{KIN}$ was selected since it is the closest non-ZRK BLAST hit to ZED1 identified in *A. thaliana*. We carried out co-IP experiments by transiently expressing each kinase with either ZAR1 or ZAR1$^{LRR}$ in *N. benthamiana*. Surprisingly, we found that PBL2 and WALK2$^{KIN}$ could interact with ZAR1 (Figure 5B) and ZAR1$^{LRR}$ (Figure 5C). BIK1 did not interact with ZAR1 or ZAR1$^{LRR}$ (Figure 2B, C, 5B), indicating that the interaction was specific. ZED1 interacted with both ZAR1 and ZAR1$^{LRR}$ (Figure 5B), as expected. The interaction between ZAR1 (or ZAR$^{LRR}$), and WALK2$^{KIN}$ or PBL2 is remarkable, as there is only 34% pairwise identity between ZED1 and WALK2$^{KIN}$, and 31% pairwise identity between ZED1 and PBL2 (Supplemental Figure S5). In addition, since ZAR1 can interact with PBL2 and WALK2$^{KIN}$ but not BIK1, this indicates that ZAR1 can achieve interaction specificity for certain kinases, and that it can interact with kinases outside of the ZRK family. Furthermore, the similarity between ZAR1-kinase and ZAR1$^{LRR}$-kinase interactions suggests that the LRR domain confers specificity and is sufficient for the kinase interactions.

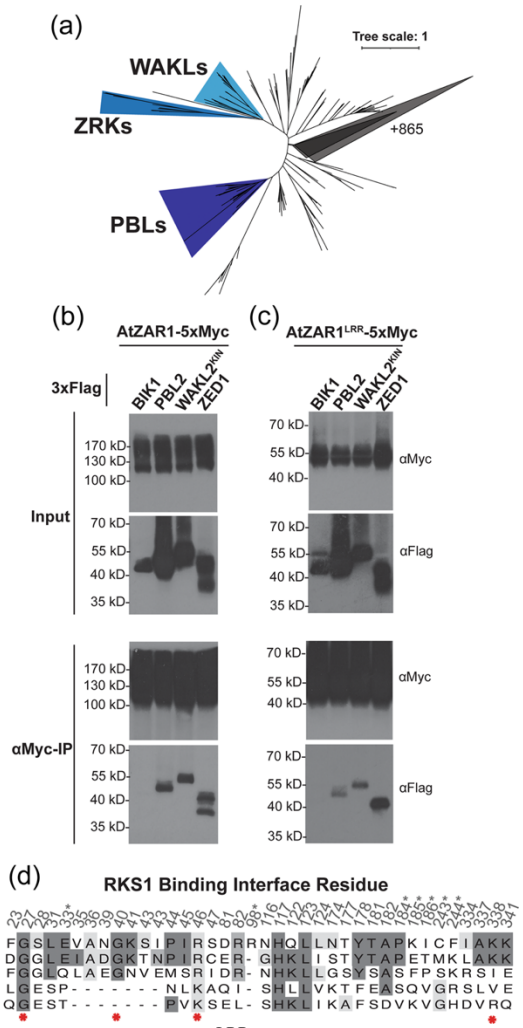



**Figure 5.** Diverse kinases interact with ZAR1 and ZAR1$^{LRR}$. (A) Phylogeny of *A. thaliana* kinases with the ZRK, PBL, and WAKL clades highlighted. Grey shading indicates collapsed branches with an additional 865 kinases. Representative western blots showing the co-immunopurification of (B) ZAR1 and (C) ZAR1$^{LRR}$ with ZED1, PBL2, and the kinase domain (aa 346-748) of WALK2 but not BOTRYTIS-INDUCED KINASE1 (BIK1, PBL family) when transiently expressed in *N. benthamiana*. Each kinase was fused with a 3xFlag tag and co-expressed with either ZAR1 or ZAR1$^{LRR}$ which were both fused with a 5xMyc tag. Anti-Myc beads were used to immunoprecipitate ZAR1 or ZAR1$^{LRR}$. All experiments were repeated three times with similar results. Predicted molecular masses of the PBL2-3xFlag and WALK2KIN-3xFlag is 49kD. (D) Amino acid alignment showing the residues and positions of the binding interface of RKS1 based on PDB:6j5w and corresponding residues in ZED1, WALK2, PBL2, and BIK1 (Wang et al., 2019a; Wang et al., 2019b). Asterisks above amino acid positions indicate a predicted ZAR1$^{NBD}$ contact. Red asterisks indicate that position was identified as under significantly stronger negative selective pressure in the ZRKs relative to the WAKLs and PBLs, based on the Hyphy Contrast-FEL model (see Supplemental Table S2).

*Selection analyses identify residues important for ZAR1 interaction specificity*

We hypothesized that residues important for ZAR1 interaction would be under significantly stronger negative selection pressure in ZRKs when compared to PBLs and WAKLs, and that ZAR1-kinase interactions outside of the ZRK family are likely not biologically relevant for pathogen recognition. While PBL2 and WALK2$^{KIN}$ can interact with ZAR1, we expected that selective pressure to gain and maintain ZAR1 interactions would be more pervasive in the ZRK family. Using the contrast-FEL model through hyphy, we identified 8 residues in the ZRKs

which were under significantly stronger negative selection compared to WAKLs and PBLs (Supplemental Figure S5; Supplemental Figure S6; Supplemental Table S2). To place this analysis in the context of ZAR1-kinase interactions, we used Chimera and the inactive RKS1-ZAR1 structure (PDB: 6J5W) to identify RKS1 residues in the ZAR1 binding interface (Figure 5D). Out of the 8 residues identified in the selection analysis, G27, G40, R46, and K338 are in the RKS1-ZAR1 binding interface (Figure 5D). As a validation of our findings, the G27A mutation in RKS1 has been previously shown to hinder ETI, while the G29E mutation in ZED1 (corresponding to G40 in RKS1) has been shown to both hinder ETI and weaken the ZAR1 interaction [24,36]. In addition, 3 positions adjacent to binding interface residues (C38, S115, E339) and one activation loop residue (K220) were under significantly stronger negative selection in the ZRKs compared to PBLs and WAKLs (Supplemental Table S2), suggesting that these residues may be functionally relevant. While PBL2 and $WALK2^{KIN}$ can interact with ZAR1, evidence from our selection analysis suggests that clear differences in selective pressure are present in the ZRK family. Moreover, the specific residues under significantly different selective pressure in the ZRK family suggest that the ZRKs are likely under selective pressure to interact with ZAR1.

*Binding interface analyses identify residues important for ZAR1 interaction specificity*

To identify additional residues that are important in the ZAR1-kinase binding landscape, we used MutaBind2 to predict the change in binding affinity ($\Delta\Delta G$) resulting from every possible single amino acid substitution in the RKS1-ZAR1 binding interface based on the previously published inactive structure (J. Wang, Hu, et al., 2019; J. Wang, Wang, et al., 2019). Mutations in RKS1 residues G27, G40, and P44 were predicted to strongly reduce binding affinity regardless of the alternative amino acid (Figure 6A). These residues stand out as binding "hotspots" and are conserved across all interacting ZRKs (Figure 6A). Both G27 and G40 were identified in our selection analysis as being under significantly stronger negative selection pressure in ZRK family when compared to WAKLs and PBLs (Figure 5D). Furthermore, both G27 and G40 have been shown to weaken the ZAR1 interaction and abolish RKS1 and ZED1-mediated ETI [24,36]. Interestingly, the $RKS1^{I43N}$ and $RKS1^{Q122K}$ mutations were predicted to strengthen RKS1 interaction, and both alternative residues are found in other interacting ZRKs (Figure 6A). We also noted that mutations in the 7 identified RKS1-$ZAR1^{NBD}$ contacts (V33, R98, P184, K185, I186, C243, and F244) were predicted to result in more neutral changes to binding affinity overall, further supporting the importance of $ZAR1^{LRR}$ in kinase binding (Figure 6A).

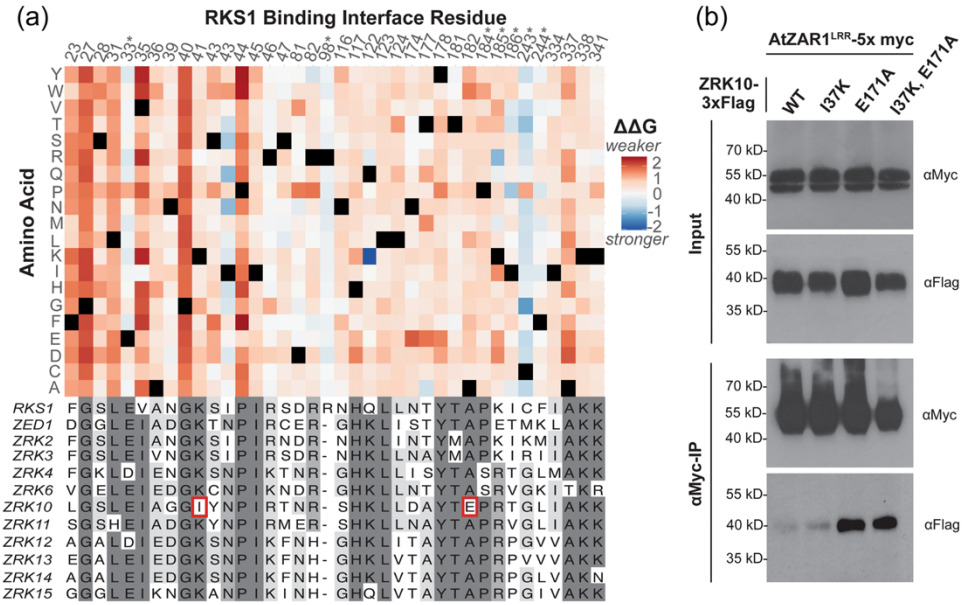



**Figure 6.** Binding interface analysis reveals RKS1 binding hotspots and residues that increase ZRK10-ZAR1$^{LRR}$ interaction strength. (A) Heat map for the predicted change in binding affinity (ΔΔG) resulting from every possible binding interface mutation in RKS1. The position of RKS1 binding interface amino acids is shown at the top of the blot, and was identified using Chimera and PDB:6j5w [24,25]. Asterisks above amino acid positions indicate a predicted ZAR1$^{NBD}$ contact. Boxes on the heatmap correspond to the predicted change in binding affinity calculated using MutaBind2 for a given single amino acid substitution on the Y-axis. Black boxes correspond to the original residue in RKS1. An amino acid alignment of RKS1 and ZAR1-interacting ZRKs is shown below the heatmap. Red boxes indicate residues targeted for mutagenesis in panel B. (B) Representative western blots showing the co-immunoprecipitation of ZAR1$^{LRR}$ with ZRK10, ZRK10$^{E171A}$ and ZRK10$^{I37K/E171A}$ in *N. benthamiana*. ZRK10 and each ZRK10 mutant were fused with a 3xFlag tag and transiently co-expressed with ZAR1$^{LRR}$ which was fused with a 5xMyc tag in *N. benthamiana*. Anti-Myc beads were used to immunoprecipitate ZAR1. Co-immunoprecipitation experiments were repeated three times with similar results.

*A single amino acid substitution in ZRK10 strengthens the interaction with ZAR1$^{LRR}$*

We sought to utilize findings from the ZAR1-kinase co-IPs and the binding interface analysis to strengthen the interaction between ZAR1$^{LRR}$ and ZRK10. We specifically focused on the ZRK10-ZAR1$^{LRR}$ interaction as it was consistently observed to be weak relative to other ZRKs (Figure 2C). We identified two residues in the RKS1 binding interface alignment that were unique to ZRK10 (Figure 6A). All ZAR1-interacting ZRKs have a lysine that aligns with RKS1$^{K41}$, while ZRK10 has an isoleucine in the corresponding position (ZRK10$^{I37}$; Figure 6A, Supplemental Figure S3). Similarly, ZAR1-interacting ZRKs have an alanine that aligns with RKS1$^{A182}$, while ZRK10 has a glutamic acid in the corresponding position (ZRK10$^{E171}$; Figure 6A, Supplemental Figure S3). In addition, the ΔΔG values from our binding interface analysis suggest that the two differences unique to ZRK10 would result in a decrease in binding affinity if they were introduced in RKS1. Specifically, RKS1$^{K41I}$, which aligns with ZRK10$^{I37}$, is predicted to weakly decrease RKS1-ZAR1 interaction strength (ΔΔG=0.52) while RKS1$^{A182E}$ (which aligns with ZRK10$^{E171}$) is predicted to result in a greater decrease in RKS1-ZAR1 interaction (ΔΔG=1.13). Since these differences in ZRK10 are predicted to negatively affect binding affinity, we investigated if introducing the consensus residue in ZRK10 would increase ZRK10-ZAR1$^{LRR}$ interaction strength.

We generated ZRK10$^{I37K}$, ZRK10$^{E171A}$ and ZRK10 $^{I37K/ E171A}$ mutants and transiently co-expressed ZAR1$^{LRR}$ fused to a 5x-Myc tag with each ZRK10 mutant as an in-frame fusion with

3x-Flag in *N. benthamiana*. We then tested for differences in interaction strength with ZAR1$^{LRR}$ using a co-IP experiment. Remarkably, we observed a consistently stronger interaction between ZRK10$^{E171A}$ and ZAR1$^{LRR}$ when compared to ZRK10 (Figure 6B). In agreement with our predictions, ZRK10$^{I37K}$ had no consistent detectable effect on the ZAR1$^{LRR}$ interaction. Similarly, we did not observe a stronger interaction between ZRK10$^{I37K/E171A}$ and ZAR1$^{LRR}$ when compared to ZRK10$^{E171A}$, suggesting that the I37K mutation had little effect on ZAR1$^{LRR}$ interaction when combined with the E171A mutation. Therefore, the binding interface analysis identified a key residue that strengthened the ZRK10-ZAR1$^{LRR}$ interaction in co-IP experiments. Moreover, we have shown that ZAR1$^{LRR}$ interaction specificity is sensitive to single amino acid substitutions.

**Discussion**

Here, we demonstrate that ZAR1 can interact with diverse *A. thaliana* kinases while simultaneously being sensitive to single amino-acid substitutions in kinase interactors. Under native conditions in *A. thaliana*, we show that ZAR1 can interact with 6 out of 13 ZRKs (Figure 1). Using an *N. benthamiana* transient overexpression system, we showed that ZAR1 can interact with all ZRKs except for ZRK7 (Figure 2). We found that some ZRKs result in a strong effector-independent HR phenotype when co-expressed with native and/or overexpressed NbZAR1 (Figure 3). Although ZAR1 can interact with an unannotated splice variant of ZRK7, subsequent analysis suggested ZRK7 is likely a pseudogene (Figure 4). While all known ZAR1-mediated effector recognition requires a kinase in the ZRK family (Laflamme et al., 2020; Lewis et al., 2013; Martel et al., 2020; Seto et al., 2017; G. Wang et al., 2015), we showed that ZAR1$^{LRR}$ can interact with kinases outside of the ZRK family with greater strength compared to some ZRKs (Figure 2, 5). The overlap between ZAR1 and ZAR1$^{LRR}$ interaction profiles demonstrates that ZAR1$^{LRR}$ is sufficient for kinase interactions and suggests that ZAR1$^{LRR}$ confers specificity for kinases (Figure 2, 5). The specificity of ZAR1$^{LRR}$ interactions can be manipulated by single amino acid substitutions in the kinase interactor (Figure 6).

Interaction profiles of ZAR1 in *A. thaliana* during HopZ1a recognition suggest that ZAR1 interacts with multiple ZRKs during effector recognition. The strongest ZAR1$^{LRR}$ interactors observed in co-IP experiments in *N. benthamiana* overlap with the ZRKs detected through LC-MS in *A. thaliana* (Figure 1, 2). This suggests that the interactions observed in *N. benthamiana* between ZAR1 and ZRK2, ZRK4, ZRK6, ZRK10, ZRK11, and ZRK12 were too weak to detect in Arabidopsis, possibly due to the sensitivity of LC-MS, experimental conditions, native expression levels, or subcellular or tissue localization of the proteins. We feel that the sensitivity of our LC-MS experiment and/or the collection of tissue samples during HopZ1a recognition is the most likely explanation for the observed differences in ZAR1-ZRK interactions between our LC-MS data in *A. thaliana* and our co-IP data in *N. benthamiana*. However, we cannot rule out the possibility that these differences may also be due to native expression in *A. thaliana* compared to overexpression in *N. benthamiana.* ZAR1 is necessary for the recognition of multiple effector proteins (Laflamme et al. 2020), and the additional ZRK interactors identified here (Supplemental Figure S7), suggest that additional effectors may be recognized through ZAR1-ZRK complexes.

A remaining question in the ZAR1 system is whether the resistosome pentamer can contain different ZRKs upon effector recognition, or if it is comprised of only the ZRK responsible for mediating recognition to a specific effector. Given the difference in ZAR1$^{LRR}$-ZRK interaction strengths, it is unclear whether all interactions are represented under native

conditions. Alternatively, if ZAR1 is present at a low level and ZRKs competitively bind to ZAR1, this would result in only identifying a subset of the strongest interacting ZRKs. We have previously shown that ZAR1 alleles from *A. thaliana* ecotypes may have stronger or weaker interaction strength with ZED1 relative to Col-0 [37]. It is possible that selection pressure imposed by pathogens can influence broader ZAR1-ZRK affinities across *A. thaliana* ecotypes. While clear differences in interaction strength were visible in $ZAR1^{LRR}$-ZRK co-IP experiments, all ZRKs appeared to interact strongly with full-length ZAR1 (Figure 2). The tomato NLR Prf forms hetero-oligomeric complexes that associate with Pto kinase and multiple Pto kinase family members in its inactive state, and these additional members contribute to *P. syringae* resistance [15]. It is tempting to speculate that ZAR1-ZRK complexes are similar to Prf-Pto kinase complexes, and that additional ZRKs other than the one required for effector recognition are present in oligomeric resistosomes. These additional ZRKs may have a role in conferring broader resistance to *P. syringae,* although further work is needed to test this hypothesis.

Our findings provide additional evidence that $ZAR1^{LRR}$ is responsible for determining guardee specificity. We did not identify any kinases that could interact with full length ZAR1 but not $ZAR1^{LRR}$ and vice-versa (Figure 2). The overlap between interaction trends with $ZAR1^{LRR}$ and ZAR1 presented here, the cryo-EM structure of ZAR1, and previous findings on the ZAR1 domains in the context of ZED1 interactions suggests that $ZAR1^{LRR}$ determines guardee specificity across all kinase interactors [6,20,24,25]. In addition, all ZRKs that interacted strongly with $ZAR1^{LRR}$ resulted in strong cell death when the ZRK and NbZAR1 were overexpressed (Figure 3). These findings further support our previous conclusion that $ZAR1^{LRR}$ interaction is required for ZAR1-mediated HR [20,37].

Using a binding interface analysis, we were able to rationally increase the interaction strength between $ZAR1^{LRR}$ and kinase interactors (Figure 6). We have previously shown that the S584R mutation in $ZAR1^{LRR}$, from *A. thaliana* ecotype Cvi-0, can strengthen the interaction between $ZAR1^{LRR}$ and ZED1 [37]. Although mutations in ZED1 can cause a loss of the $ZAR1^{LRR}$ interaction [20], mutations in a ZAR1 guardee that increase affinity for $ZAR1^{LRR}$ have not been previously described. Here, we show that a single amino acid substitution in ZRK10 was sufficient to increase interaction strength with $ZAR1^{LRR}$ (Figure 6). Our selection of $ZRK10^{E171A}$ was based on the consensus residue of ZRKs that interact with ZAR1 at that position, as well as the predicted effect of mutating that site based on the RKS1-ZAR1 binding interface analysis. It is possible that our rationale for targeting this specific residue can be extended to improve ZAR1 interaction strength with other kinases. Engineering NLRs is a complex problem since it requires balancing protein-protein interactions with appropriate activation of the NLR. The NLR RPS5 has been engineered to enable recognition of different effectors. RPS5 is normally activated when its kinase guardee PBS1 is cleaved by *P. syringae* effector AvrPphB (Ade et al. 2007). Kim and colleagues (2016) were able to introduce new cleavage sites into PBS1, allowing recognition of proteases from other pathogens [38]. More recently, the recognition specificity of the rice NLR RGA5 was expanded through modifications of the RGA5 integrated-domain effector-binding interface, leading to recognition of an additional effector [39]. The ability to fine-tune ZAR1-kinase interaction strength is a major advancement towards the rational design of a ZAR1 guardee based on existing pathogen-targeted kinases that are not guarded by ZAR1.

The *A. thaliana* Col-0 genome contains a version of ZRK7 which can interact with ZAR1 and $ZAR1^{LRR}$, while the annotated version of ZRK7 is not capable of ZAR1 or $ZAR1^{LRR}$ interaction (Figure 2, 4). We were unable to find concrete evidence that *gZRK7* is transcribed in its entirety or is functionally relevant for ZAR1 immunity. However, we did find evidence of

intron retention in the annotated *ZRK7* gene based on publicly available RNA-seq data. Furthermore, evidence of *ZRK7* intron retention was previously identified in transcriptomic analysis investigating alternative splicing in *A. thaliana* [35]. Several ZRKs, including *ZRK7*, are present in a genomic cluster in the *A. thaliana* genome (Figure 2A) [6], suggesting tandem duplication as a mechanism for generating ZRK diversity. Pseudogenes are frequently derived from tandem gene-duplication events in *A. thaliana*, and may represent an intermediate step in gene duplication when found in this context [40]. A previous genome-wide analysis of *A. thaliana* pseudogenes identified ZRK7 as a pseudogene based on two separate prediction pipelines [41]. ZRK7 also has high sequence similarity to a known *A. thaliana* pseudogene At1g43895 [41]. It appears that *ZRK7* is a pseudogene and might have the potential to functionalize under the right selective pressure. It is also possible that genomic ZRK7 may have a detrimental effect (i.e., cell death) when expressed in *A. thaliana*, and became a pseudogene to avoid this. Tandemly arrayed genes involved in biotic stress are often under strong selection pressure, and can result in pseudogenes [42]. Our selection analysis did identify selective pressure unique to amino acids at the N-terminus of gZRK7, as well as gene level positive selection on *gZRK7, ZRK12, ZRK13*, and *ZRK15* (Supplemental Figure S4, Supplemental Table S1). Further investigation of ZRK7 may yield insight into the evolution of *A. thaliana* ZRKs.

We have shown that interspecific pairings of ZAR1 orthologs and ZRKs can lead to an autoactive HR. We found that some ZRKs results in a strong effector-independent HR phenotype when co-expressed with native and/or overexpressed NbZAR1 (Figure 3). While others have observed an HR resulting from overexpressing RKS1 with several different ZAR1 orthologs including NbZAR1 [22], our data highlights the differential levels of autoactivity resulting from one gene family paired with one ZAR1 ortholog. Heterologous expression of certain NLRs has been previously shown to result in cell death in the absence of pathogen effectors due to the lack of negative regulation from a "paired" NLR or guardee [43–45]. Here we show that heterologous expression of a guardee results in autoactive cell death mediated by an endogenous NLR. There are several possible explanations for the interspecific ZAR1-ZRK autoactive HR. First, NbZAR1 may be easier to activate than AtZAR1. In support of this hypothesis, NbZAR1 can be autoactivated by introducing the HRD mutation, while the same mutations do not result in an autoactive AtZAR1 protein [23,46]. Second, interspecific ZRK and PBL pairings may result in autoactivation of NbZAR1 immunity, due to a lack of co-evolution between these gene families. While the ZRK-PBL co-evolution hypothesis is possible, we think it is unlikely since this same mismatch between ZRKs and PBLs should also activate AtZAR1 which was not consistently observed (Supplemental Figure S2). Third, the interaction between NbZAR1 and ZRKs may result in a clash with the NbZAR1$^{NBD}$. Based on the cryo-EM structure, AtZAR1 activation occurs after the RKS1 activation loop is stabilized by PBL2 interaction [24,25]. In the inactive state, the activation loop of RKS1 is not well resolved. However, several other regions of RKS1 are near the AtZAR1$^{NBD}$. Slight differences in binding orientation between the cryo-EM structure and the ZRK-NbZAR1 interaction could result in a clash between ZRKs and NbZAR1$^{NBD}$, leading to activation in the absence of effector perturbation (J. Wang, Hu, et al., 2019b; J. Wang, Wang, et al., 2019b). Identifying the cause of interspecific autoactive HR phenotypes is a crucial next step towards interspecific pairings of *A. thaliana* ZRK recognition pathways with ZAR1 orthologs in non-model species.

The current mechanistic model of ZAR1 activation based on the cryo-EM structure is unable to connect some experimental findings within the ZAR1 system, and many unanswered questions about ZAR1 function under native conditions remain. Previous studies have

demonstrated that ZRKs may interact with diverse PBLs in a dynamic fashion in the presence and absence of pathogen effectors [10,47]. We observed a ZAR1-PBL2 interaction under transient overexpression in a heterologous system in the absence of a pathogen effector (Figure 5). This was not observed in the ZAR1-RKS1-PBL2 structure or previous work on AvrAC recognition in Arabidopsis [7,24,25]. However, we previously demonstrated that the *N. benthamiana* system is remarkably sensitive and can distinguish weak interactions that may result in loss-of-function in Arabidopsis (Baudin et al. 2017, Baudin et al. 2021). While it is possible that PBL2 interacts with ZAR1 through an endogenous ZRK in *N. benthamiana*, we find this unlikely since previous studies have shown that PBL2 does not interact with RKS1 in the absence of the pathogen effector [7]. ZAR1 interaction specificity under native conditions may be limited to the strongest ZRK interactors due to competitive binding, preventing ZAR1-kinase interactions outside of the ZRK family. In addition, ZRKs may have different tissue or developmental expression compared to PBL2. We showed that ZAR1 also displays strong interactions with WALK2$^{KIN}$, which is the closest non-ZRK kinase to ZED1. The ZRKs are believed to have evolved from WAKs through the loss of their extracellular domain [22], which would affect their subcellular localization (See Appendix I). The subsequent expansion of ZRKs coupled with this change in localization would give ZAR1 the opportunity for selective and versatile interactions with diverse kinases, resulting in the current immunodiversity conveyed by this key plant immunity hub.

**Materials and Methods**

*Molecular cloning*

All gene fragments used in cloning were amplified with Phusion polymerase (New England Biolabs), and all constructs were confirmed by sequencing. All *ZRK*s, in addition to *PBL2*, and *WALK2* were amplified with a 5'*Xho*I site. The pMac16 vector was modified from pBD to contain a 3xFlag tag between the *Stu*I and *Spe*I sites, and has been previously described [23]. The pMac16 vector was digested with *Xho*I and *Stu*I to insert *ZRK*s, *WAKL2*, and *PBL2*. *BIK1* was amplified to contain attB1 and attB2 sites, then cloned into a pDonor207 vector through a BP reaction. We generated a gateway compatible pMac16 vector (hereafter, pMac16-GW) using the gateway conversion kit (Invitrogen) to insert a gateway cassette into *Stu*I digested pMac16. We then used the LR reaction to clone *BIK1* into this vector. The *ZED1*, *AtZAR1*, *AtZAR1$^{LRR}$*, *NbZAR1* and virus-induced gene silencing (VIGS) constructs were described in previous studies [20,37]. The dexamethasone-inducible *hopZ1a* construct was described previously [48]. For site directed mutagenesis, *ZRK10* was amplified to contain attB1 and attB2 and cloned into pDonor207 using the BP reaction. Mutations were introduced using the Q5 site directed mutagenesis kit (New England Biolabs). *ZRK10* point mutants were cloned into pMac16-GW using the LR reaction. The *ZAR1* constructs used for the genetic transformation of *A. thaliana* were designed by amplifying a genomic fragment of *ZAR1* containing 1 kb of the promoter region and the coding sequence, or the coding sequence and a 3xFlag tag [37]. The fragments were amplified by PCR to add a 5' *Sal*I site and an attB1 site, and a 3' *Not*I site and an attB2 site. The resulting *promoter-ZAR1-3xFlag* or *promoter-ZAR1* products were first cloned into pBS in a three-body ligation, followed by a BP reaction into pDONR207 and an LR reaction into linearized pEarleyGate303 [49].

*Plant material and growth conditions*

*N. benthamiana* and *A. thaliana* plants were grown as previously described [20]. *N. benthamiana* seeds were sterilized in 10% bleach and germinated on soil. *A. thaliana* seeds were

vapor sterilized and germinated on 0.5x Murashige and Skoog 0.8% agar media. *N. benthamiana* seedlings were transplanted on Sunshine #1 (Sun Gro Horticulture) soil supplemented with 14:14:14 Osmocote and grown in a growth chamber at 22°C with 9 hours of light (~130 $\mu Em^{-2} s^{-1}$). For *A. thaliana*, Sunshine #1 soil (lacking RSi) was used and supplemented with 20:20:20 fertilizer [50].

*Generation and validation of transgenic lines*

*A. thaliana zar1-1* mutants [5] were transformed with native promoter-driven *ZAR1* carrying an in-frame 3xFlag epitope tag or a native promoter-driven *ZAR1* construct lacking the 3xFlag tag. Transformations were conducted via the floral dip method using *A. tumefaciens* GV3101. T1 individuals were selected by Basta resistance and confirmed to be *zar1-1* by PCR. Candidate T1 lines were screened for restoration of the HR when plants were pressure-infiltrated with *P. syringae* pathovar tomato DC3000 carrying *hopZ1a* [5]. The number of insertions in the T2 generation was determined based on the percentage of plants that survived growth on 0.5x MS containing 6mg/L bialophos, and T2 lines with one insertion were selected. Homozygous lines in the T3 generation were identified based on the percentage of plants surviving on 0.5x MS containing 6mg/L bialophos. Protein expression in T3 lines was checked via anti-Flag western blot.

*N. benthamiana transient expression, virus-induced gene silencing, and hypersensitive response assays*

For transient gene expression, *A. tumefaciens* GV2260 was grown overnight at 28°C in LB containing appropriate antibiotics. Bacteria were resuspended in induction media (10 mM MES pH 5.6 and 200 μM of acetosyringone) and incubated at room temperature in the dark for ~4 hours. The optical density at 600nm ($OD_{600}$) of each culture was adjusted to 0.25. The leaves of 4–7-week *N. benthamiana* plants were infiltrated with bacteria using a needle-less syringe. 24 hours post-infiltration, 40 μM dexamethasone was sprayed on leaves. To assay for the hypersensitive response, plants were scored 24 hours after dexamethasone induction based on 4 categories (Figure 3A). The scores were defined as no HR, weak HR (<25% of leaf area), medium HR (<75% of leaf area) or strong HR (75-100% of leaf area). *N. benthamiana* plants silenced for *NbZAR1* or *GUS* were prepared as previously described [20]. Briefly, *A. tumefaciens* was prepared as described for transient assays. The 5$^{th}$ and 6$^{th}$ leaves of 3-week-old *N. benthamiana* plant were infiltrated completely with bacteria using a needle-less syringe. Plants were used 3 weeks after initial VIGS inoculation.

*Liquid chromatography tandem mass spectroscopy*

For each *pZAR1:ZAR1-3xFlag* or *pZAR1:ZAR1* line, 6- to 7-week-old plants were sprayed with *Pto*DC3000 carrying *hopZ1a*. Bacteria for spray inoculation were resuspended in 10mM $MgCl_2$ with 0.02% Silwet L-77, adjusted to $OD_{600}$ of 0.8, and sprayed until leaves were wet, as previously described [50]. The total protein was extracted from 3 g of leaf tissue harvested 3 days after spray inoculation, and the mixture was incubated with magnetic beads coated with anti-Flag antibodies. The beads were washed 3 times with low salt buffer (50 mM HEPES, 50 mM NaCl, 10 mM ethylene diamine tetraacetic acid (EDTA), pH 7.5), 3 times with high salt buffer (50 mM HEPES, 150 mM NaCl, 10 mM EDTA, pH 7.5) and twice with phosphate buffer on a magnetic stand. The low salt and high salt buffers were supplemented with a homemade proteinase inhibitor cocktail (1 mM PMSF, 1 mg/ml leupeptin [Sigma-Aldrich L2023], 1 mg/ml

aprotinin [Sigma-Aldrich A6191], 1 mg/ml antipain [Sigma-Aldrich A1153], 1 mg/ml chymostatin [Sigma-Aldrich C7268] and 1 mg/ml pepstatin [Sigma-Aldrich P5315]). The beads were resuspended in 20 µL of ammonium bicarbonate buffer with 5% of acetonitrile. Next, 10 µL of 15 mM DTT was added and the reaction was incubated for 1 hour at 37°C, followed by the addition of 40 µL of 22 mM iodoacetamide and another 1 hour incubation in the dark. Lastly, 20 µL of 22 mM DTT and 10 µL of 100 ng/µL trypsin (Promega) were added, and the reaction was incubated overnight at 37°C. The tryptic digest was centrifuged to pellet the magnetic beads. The peptide-containing supernatant was removed and placed into a 96-well plate.

The peptides were analyzed using a Thermo Scientific, Orbitrap Elite mass spectrometer (Thermo Scientific, Waltham, MA). Proteins were identified with Mascot Server (Matrix Science Inc., Boston, MA) while Scaffold (version Scaffold_4.5.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95% probability. For additional details on mass spectrometry and data analysis, see supplemental materials and methods.

*Co-immunoprecipitation experiments*

Tissue samples were harvested from at least 2 separate *N. benthamiana* plants 48 hours after *A. tumefaciens* infiltration and 24 hours after dexamethasone induction. For co-IP experiments, 3 $cm^2$ of tissue was collected. For experiments involving RKS1 and ZRK13, *NbZAR1*-VIGS plants were used to avoid cell death.

Tissue samples were ground using liquid nitrogen and homogenized in 1 mL IP1 buffer (50 mM HEPES, 50 mM NaCl, 10 mM ethylene diamine tetraacetic acid (EDTA), 0.2% Triton X-100 and 0.1 mg $ml^{-1}$ Dextran [Sigma-Aldrich D1037, St. Louis, MO], pH 7.5). Samples were incubated with 5 µl Myc-Trap magnetic beads for 2.5 hours on a spinning wheel at 4°C. Beads were subjected to four quick washes (tubes inverted 6 times) followed by two long washes (five minutes on spinning wheel) with IP2 (50 mM HEPES, 250 mM NaCl for $ZAR1^{LRR}$ or 450mM NaCl for ZAR1, 10 mM EDTA and 0.1% Triton X-100, pH 7.5). Both IP1 and IP2 were supplemented with the same homemade protease inhibitor cocktail as for the mass spectrometry experiment. The beads and input fraction were boiled with Laemmli buffer, and each fraction was separated on a 10% SDS-PAGE gel. Proteins were detected using an anti-Myc antibody (Abcam ab32072, Cambridge, MA), or anti-Flag antibody (Sigma- Aldrich F1804) followed by a secondary antibody coupled with horseradish peroxidase.

*Binding interface, phylogenetic and selection analysis*

Chimera was utilized to identify contacts between RKS1 and ZAR1 based on the inactive ZAR1-RKS1 cryo-EM structure (PDB:6J5W; Pettersen et al., 2004; J. Wang, Hu, et al., 2019a; J. Wang, Wang, et al., 2019). MutaBind2 was used to predict the change in binding affinity (ΔΔG) resulting from every possible mutation in RKS1 in the inactive ZAR1-RKS1 cryo-EM structure [52]. Results were converted into a heatmap using the ggplot2 in R [53,54]. To test for selection, kinase protein sequences were aligned using MUSCLE and columns with gaps were removed with Belvu [55,56]. To generate a codon alignment, nucleotide sequences were aligned to the corresponding protein alignments using PAL2NAL [57]. The Datamonkey web interface was used to test for selection within codon alignments using adaptive Branch Site REL (aBSREL), and Contrast Fixed Effects Likelihood (Contrast-FEL) Hyphy models [58–61].

The *A. thaliana* kinase phylogeny was generated by searching the Arabidopsis proteome in Uniprot using the Pfam protein kinase Hidden Markov Model (PF00069). Significant matches were aligned using MUSCLE, and columns and sequences with excessive gaps were removed using Belvu [55,56]. A Maximum Likelihood Tree was constructed using RAxML and the tree was viewed in iTOL [62,63].

**References**


1. Ngou, B. P. M., Ahn, H. K., Ding, P. & Jones, J. D. G. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature* **592**, 110–115 (2021).
2. Yuan, M. *et al.* Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature* **592**, 105–109 (2021).
3. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
4. Jones, J. D. G., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, 1117 (2016).
5. Lewis, J. D., Wu, R., Guttman, D. S. & Desveaux, D. Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. *PLoS Genet.* **6**, e1000894 (2010).
6. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the Pseudomonas syringae type III effector HopZ1a. *Proc. Natl. Acad. Sci. USA* **110**, 18722–18727 (2013).
7. Wang, G. *et al.* The decoy substrate of a pathogen effector and a pseudokinase specify pathogen-induced modified-self recognition and immunity in plants. *Cell Host Microbe* **18**, 285–295 (2015).
8. Seto, D. *et al.* Expanded type III effector recognition by the ZAR1 NLR protein using ZED1-related kinases. *Nat. Plants* **3**, 25–28 (2017).
9. Martel, A. *et al.* Immunodiversity of the Arabidopsis ZAR1 NLR Is conveyed by receptor-like cytoplasmic kinase sensors. *Front. Plant Sci.* **11**, 1290 (2020).
10. Seto, D. *et al.* The small molecule Zaractin activates ZAR1-mediated immunity in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **118**, 1–8 (2021).
11. Chung, E. H. *et al.* Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. *Cell Host Microbe* **9**, 125–136 (2011).
12. Liu, J., Elmore, J. M., Lin, Z.-J. D. & Coaker, G. A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, leading to the activation of a plant innate immune receptor. *Cell Host Microbe* **9**, 137–146 (2011).
13. Axtell, M. J. & Staskawicz, B. J. Initiation of RPS2-specified disease resistance in Arabidopsis Is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**, 369–377 (2003).
14. Mackey, D., Holt, B. F., Wiig, A. & Dangl, J. L. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743–754 (2002).
15. Gutierrez, J. R. *et al.* Prf immune complexes of tomato are oligomeric and contain multiple Pto-like kinases that diversify effector recognition. *Plant J.* **61**, 507–518 (2010).
16. Ade, J., DeYoung, B. J., Golstein, C. & Innes, R. W. Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc. Natl. Acad. Sci. USA* **104**, 2531–2536 (2007).

17. Martel, A. *et al.* Metaeffector interactions modulate the type III effector-triggered immunity load of Pseudomonas syringae. *PLoS Pathog.* **18**, e1010541 (2021).
18. Zheng, X. *et al.* Two plant NLR proteins confer strain-specific resistance conditioned by an effector from Pseudomonas syringae pv. actinidiae. *J. Genet. Genomics* **49**, 823–832 (2022).
19. Breit-McNally, C., Laflamme, B., Singh, R. A., Desveaux, D. & Guttman, D. S. ZAR1: Guardian of plant kinases. *Front. Plant Sci.* **13**, 1–8 (2022).
20. Baudin, M., Hassan, J. A., Schreiber, K. J. & Lewis, J. D. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol.* **174**, 2038–2053 (2017).
21. Schultink, A., Qi, T., Bally, J. & Staskawicz, B. Using forward genetics in Nicotiana benthamiana to uncover the immune signaling pathway mediating recognition of the Xanthomonas perforans effector XopJ4. *New Phytol.* **221**, 1001–1009 (2019).
22. Gong, Z. *et al.* The origin and evolution of a plant resistosome. *Plant Cell* **34**, 1600–1620 (2022).
23. Baudin, M., Schreiber, K. J., Martin, E. C., Petrescu, A. J. & Lewis, J. D. Structure–function analysis of ZAR1 immune receptor reveals key molecular interactions for activity. *Plant J.* **101**, 352–370 (2019).
24. Wang, J. *et al.* Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science* **364**, eaav5868 (2019).
25. Wang, J. *et al.* Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science* **364**, eaav5870 (2019).
26. Shao, F. *et al.* Cleavage of Arabidopsis PBS1 by a bacterial type III effector. *Science* **301**, 1230–1233 (2003).
27. Rufián, J. S. *et al.* The bacterial effector HopZ1a acetylates MKK7 to suppress plant immunity. *New Phytol.* **231**, 1138–1156 (2021).
28. Zhang, J. *et al.* Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a Pseudomonas syringae effector. *Cell Host Microbe* **7**, 290–301 (2010).
29. Schreiber, K. J., Chau-Ly, I. J. & Lewis, J. D. What the wild things do: Mechanisms of plant host manipulation by bacterial type III-secreted effector proteins. *Microorganisms* **9**, 1–48 (2021).
30. Lehti-Shiu, M. D. & Shiu, S.-H. S. H. Diversity, classification and function of the plant protein kinase superfamily. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2619–2639 (2012).
31. Lewis, J. D., Lo, T., Bastedo, P., Guttman, D. S. & Desveaux, D. The rise of the undead: Pseudokinases as mediators of effector-triggered immunity. *Plant Signal. Behav.* **9**, e27563 (2014).
32. Laflamme, B. *et al.* The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. *Science* **367**, 763–768 (2020).
33. Berardini, T. Z. *et al.* The arabidopsis information resource: Making and mining the 'gold standard' annotated reference plant genome. *Genesis* **53**, 474–485 (2015).
34. Lamesch, P. *et al.* The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40**, 1202–1210 (2012).
35. Marquez, Y., Brown, J. W. S., Simpson, C., Barta, A. & Kalyna, M. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.* **22**, 1184–1195 (2012).

36. Hu, M., Qi, J., Bi, G. & Zhou, J.-M. Bacterial effectors induce oligomerization of immune receptor ZAR1 in vivo. *Mol. Plant* **13**, 793–801 (2020).
37. Baudin, M. *et al.* A natural diversity screen in Arabidopsis thaliana reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex. *Plant. Cell Environ.* **44**, 629–644 (2021).
38. Kim, S. H., Qi, D., Ashfield, T., Helm, M. & Innes, R. W. Using decoys to expand the recognition specificity of a plant disease resistance protein. *Science* **351**, 684–687 (2016).
39. Césari, S. *et al.* New recognition specificity in a plant immune receptor by molecular engineering of its integrated domain. *Nat. Commun.* **13**, 1–13 (2022).
40. Mascagni, F., Usai, G., Cavallini, A. & Porceddu, A. Structural characterization and duplication modes of pseudogenes in plants. *Sci. Rep.* **11**, 1–14 (2021).
41. Xiao, J. *et al.* Pseudogenes and their genome-wide prediction in plants. *Int. J. Mol. Sci.* **17**, 1991 (2016).
42. Zou, C. *et al.* Evolutionary and expression signatures of pseudogenes in Arabidopsis and rice. *Plant Physiol.* **151**, 3–15 (2009).
43. Day, B. *et al.* Molecular basis for the RIN4 negative regulation of RPS2 disease resistance. *Plant Cell* **17**, 1292–1305 (2005).
44. Césari, S. *et al.* The NB - LRR proteins RGA 4 and RGA 5 interact functionally and physically to confer disease resistance. *EMBO J.* **33**, 1941–1959 (2014).
45. Huh, S. U. *et al.* Protein-protein interactions in the RPS4/RRS1 immune receptor complex. *PLoS Pathog.* **13**, 1–22 (2017).
46. Harant, A., Pai, H., Sakai, T., Kamoun, S. & Adachi, H. A vector system for fast-forward studies of the HopZ-Activated Resistance1 (ZAR1) resistosome in the model plant Nicotiana benthamiana. *Plant Physiol.* **188**, 70–80 (2022).
47. Bastedo, D. P. *et al.* Perturbations of the ZED1 pseudokinase activate plant immunity. *PLoS Pathog.* **15**, e1007900 (2019).
48. Lewis, J. D., Abada, W., Ma, W., Guttman, D. S. & Desveaux, D. The HopZ family of Pseudomonas syringae type III effectors require myristoylation for virulence and avirulence functions in Arabidopsis thaliana. *J. Bacteriol.* **190**, 2880–2891 (2008).
49. Earley, K. W. *et al.* Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616–629 (2006).
50. Hassan, J. A., de la Torre-Roche, R., White, J. C. & Lewis, J. D. Soil mixture composition alters Arabidopsis susceptibility to Pseudomonas syringae infection. *Plant Direct* **2**, 1–13 (2018).
51. Pettersen, E. F. *et al.* UCSF Chimera-A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
52. Zhang, N. *et al.* MutaBind2: predicting the impacts of single and multiple mutations on protein-protein interactions. *iScience* **23**, 100939 (2020).
53. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. (2016).
54. R Core Team. R: A language and environment for statistical computing. (2019).
55. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
56. Sonnhammer, E. L. L. & Hollich, V. Scoredist: A simple and robust protein sequence distance estimator. *BMC Bioinformatics* **6**, 1–8 (2005).
57. Suyama, M., Torrents, D. & Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609–W612

(2006).
58. Weaver, S. *et al.* Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Mol. Biol. Evol.* **35**, 773–777 (2018).
59. Kosakovsky Pond, S. L. *et al.* HyPhy 2.5 - a customizable platform for evolutionary hypothesis testing using phylogenies. *Mol. Biol. Evol.* **37**, 295–299 (2020).
60. Smith, M. D. *et al.* Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
61. Kosakovsky Pond, S. L. & Frost, S. D. W. Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208–1222 (2005).
62. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
63. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
64. Keller, A. D., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of protein identifications made by MS/MS and database search. *Anal. Chem.* **74**, 37–38 (2002).
65. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658 (2003).
66. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, 320–324 (2014).

**Supplemental Materials and Methods**

*Liquid chromatography tandem mass spectroscopy*

The plate was placed in the autosampler of an Eksigent nanoLC 400 interfaced to an Orbitrap Elite mass spectrometer with a PicoChip nanospray source (New Objective, Woburn, MA). For each LC−MS run, an 8 μL portion of sample was loaded by the autosampler onto a 75 μm ID column containing 10 cm of 3 μm, 120 Å, ReproSil-Pur C18-AQ reverse phase packing (New Objective, Woburn, MA). Samples were eluted into the mass spectrometer with a binary gradient flow at a rate of 300 nL/min. Solvent A was 2% acetonitrile, 0.1% formic acid in water, and Solvent B was acetonitrile (Optima LC/MS grade containing 0.1% formic acid; Fisher Scientific, Waltham, MA). The gradient was programmed from 5% Solvent B to 35% Solvent B over 45 min, held at 35% B for 13 min, then returned to 5% B over 2 min. Peptides were detected in the Orbitrap with the FT survey scan from 400 to 2000 m/z at a resolution of 60,000. The 10 most intense peaks above a threshold of 30,000 counts were subjected to collision-induced dissociation (CID). The normalized collision energy was set to 35, default charge state set to 2, isolation width set to 2.0 m/z, and activation time set to 30 ms. Product spectra were recorded at 30000 ppm resolution with the low mass set to 50 Da. Dynamic exclusion was enabled for a duration of 6 seconds with a repeat count of 1. Charge state screening allowed +2 and greater charge states to be selected for CID fragmentation. Monoisotopic precursor selection was enabled.

Tandem mass spectra were extracted from Thermo RAW files using msconvert software, (Proteowizard version 3.0.9393; Palo Alto, CA; https://proteowizard.sourceforge.io/). Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot Server (Matrix Science, London, UK; version 2.5.1) and X! Tandem (The GPM, thegpm.org; version CYCLONE (2010.12.01.1)). Mascot Server was set up to search the At_cRAP-REV-CONCAT-160321 database (derived from NCBI database on 2010-09-18, along with common contaminants and protein tags/flags, reversed and concatenated on 2016-03-21; 195,554 entries) assuming the digestion enzyme trypsin. X! Tandem was set up to iteratively search subsets of the At_cRAP-REV-CONCAT-160321 database, also assuming trypsin. Mascot and X! Tandem were searched with a fragment ion mass tolerance of 0.60 Da and a parent ion tolerance of 0.60 Da. Carbamidomethylation of cysteine was specified in Mascot and X! Tandem as a fixed modification. Dehydrated of the n-terminus, glu->pyro-Glu of the n-terminus, ammonia-loss of the n-terminus, gln->pyro-Glu of the n-terminus, deamidated of asparagine and glutamine and oxidation of methionine were specified in X! Tandem as variable modifications. Oxidation of methionine was specified in Mascot as a variable modification.

Peptide Probabilities from Mascot were assigned by the Scaffold Local FDR algorithm. Peptide Probabilities from X! Tandem were assigned by the Peptide Prophet algorithm [64] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.9% probability and contained at least 3 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [65]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Spectrum counting was used as a label-free quantitation method using Scaffold protein identification software (version 4.5.1). The number of fragment spectra identifying peptides of a given protein are counted and compared.

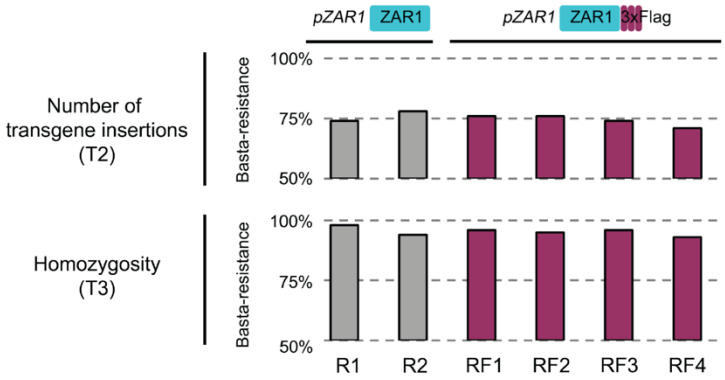



**Figure S1.** Identification of single insertion T2 and homozygous T3 *A. thaliana* stable transgenic lines. *A. thaliana zar1-1* mutants were transformed with native promoter-driven *ZAR1* construct carrying an-frame 3xFlag epitope tag or a native promoter driven ZAR1 construct lacking the 3xFlag tag. Transformations were conducted via the floral dip method using *A. tumefecians* GV3101. T1 individuals were selected by Basta resistance and confirmed by PCR that the plants were *zar1-1*. The number of insertions in the T2 generation was determined based on the percentage of plants that survived Basta selection. Homozygous lines in the T3 generation were identified based on the percentage of plants surviving Basta selection.

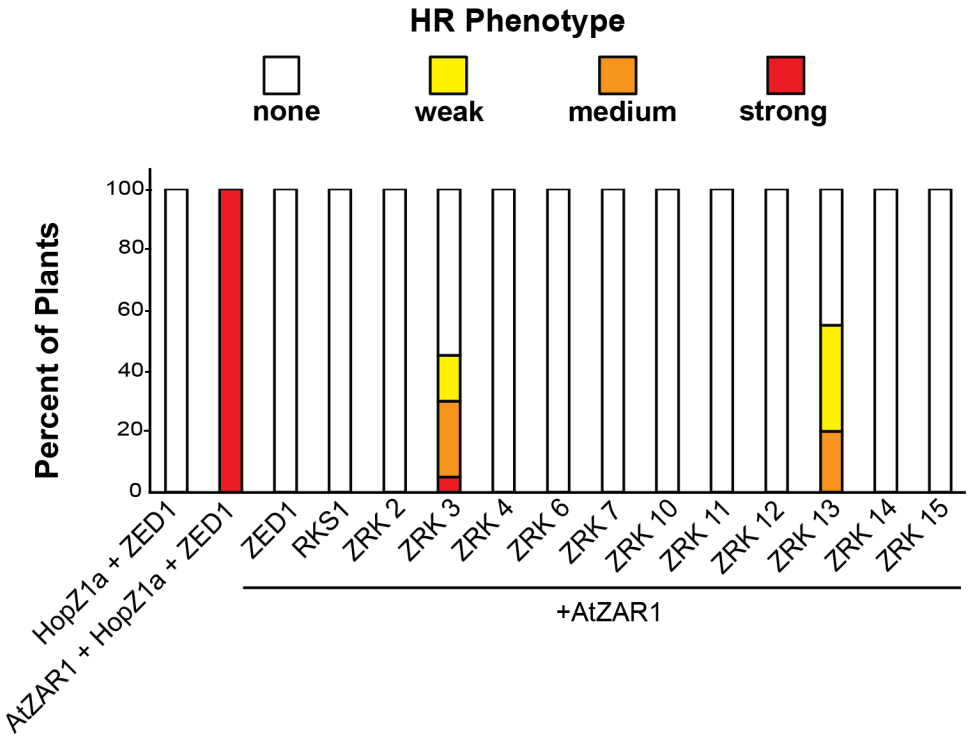



**Figure S2.** Overexpressing AtZAR1 with ZRKs does not result in consistent or strong autoactive cell death. Each ZRK was transiently overexpressed with *AtZAR1* in *N. benthamiana* plants silenced for *NbZAR1* using viral-induced gene silencing. Gene expression was induced 24 hrs following *A. tumefaciens* infiltration and macroscopic HR was scored on a 4-point scale 24 hrs after induction. For all HR data presented, 20 plants were used across 3 separate experiments.




**Figure S3.** Amino acid alignment of the ZRK family. Protein sequences were aligned using MUSCLE and visualized with ESPrint3x [55,66].

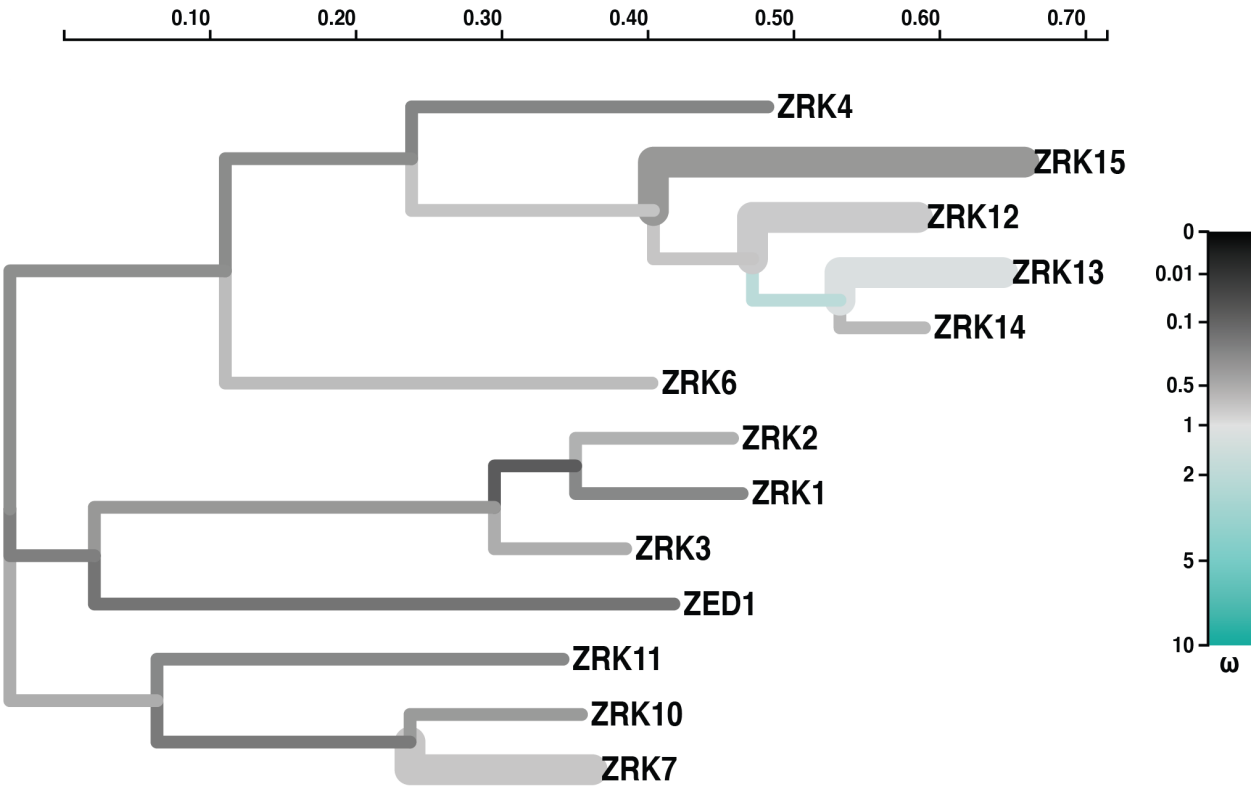



**Figure S4.** Phylogenetic analysis finds evidence of diversifying selection acting on ZRK7, ZRK12, ZRK13 and ZRK15. Kinase sequences were aligned using MUSCLE and columns with gaps were removed with Belvu [55,56]. To generate a codon alignment, nucleotide sequences were aligned to the corresponding protein alignments using PAL2NAL [57]. The Datamonkey web interface was used to test for selection within codon alignments using the adaptive Branch Site REL (aBSREL) Hyphy model [58–61].

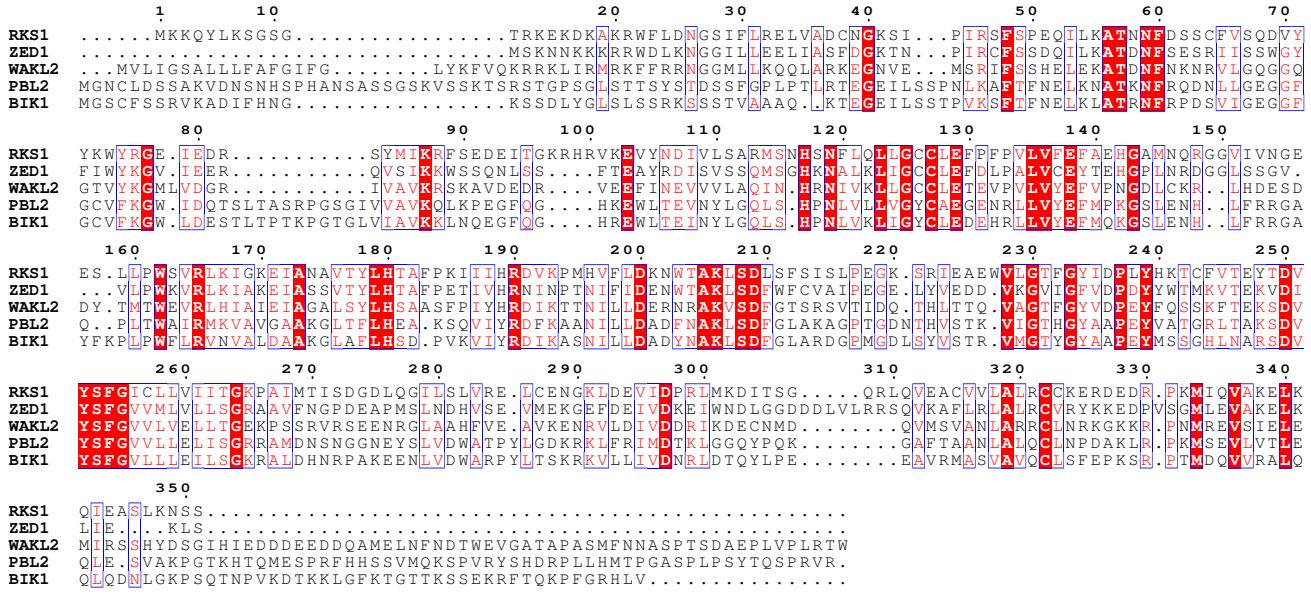


**Figure S5.** Amino acid alignment of ZED1, PBL2, WALK2$^{KIN}$, and BIK1. Protein sequences were aligned using MUSCLE and visualized with ESPrint3x [55,66].

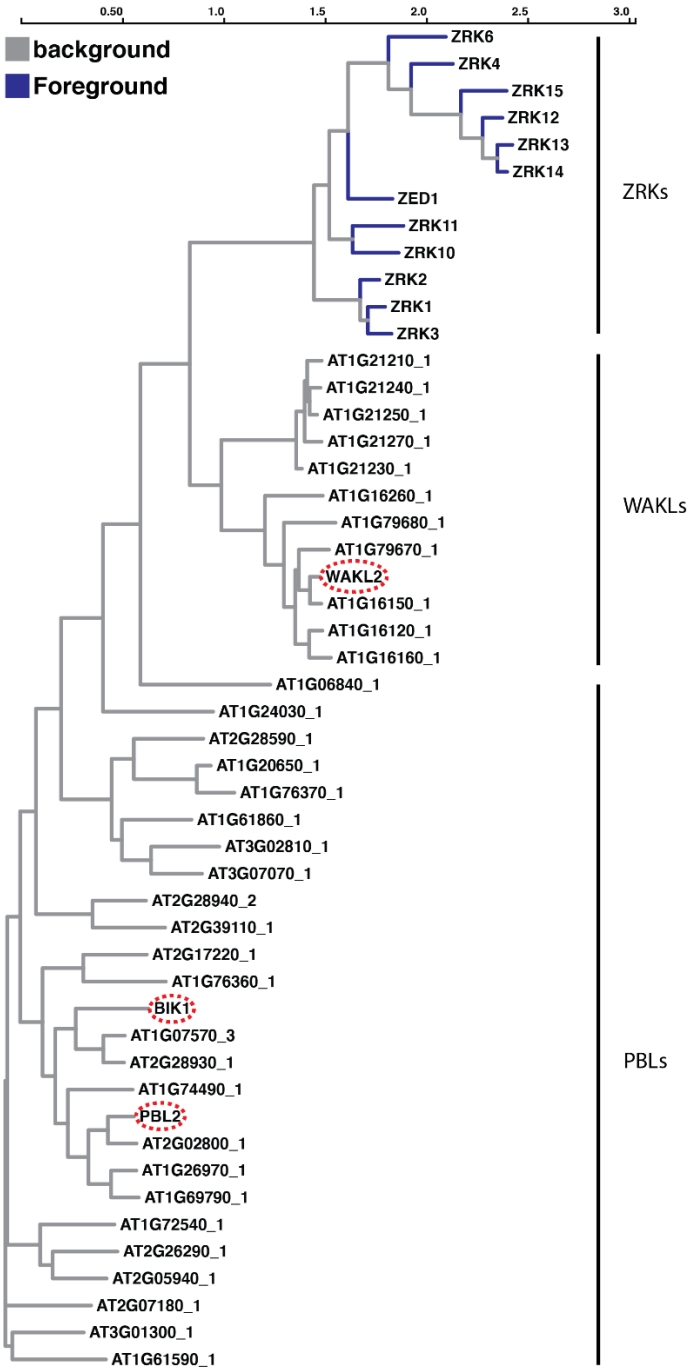



**Figure S6.** Phylogeny of ZRKs, WAKLs, and PBLs used for Hyphy Contrast-FEL analysis. Kinase sequences were aligned using MUSCLE and columns with gaps were removed with Belvu [55,56]. To generate a codon alignment, nucleotide sequences were aligned to the corresponding protein alignments using PAL2NAL [57]. The Datamonkey web interface was used to test for selection within codon alignments using the Contrast Fixed Effects Likelihood (Contrast-FEL) Hyphy models, with the ZRKs selected as the foreground [58–61].

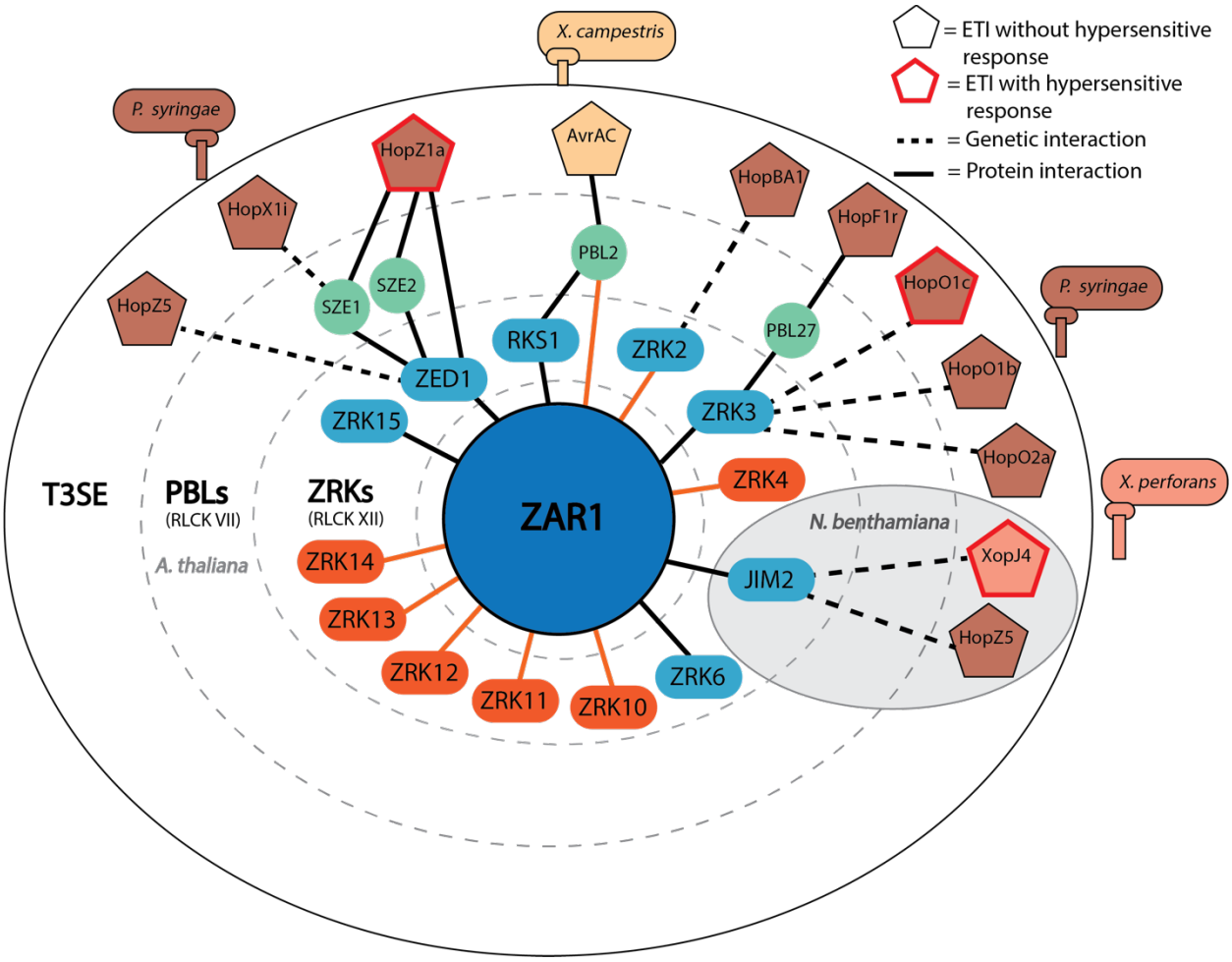



**Figure S7**. The current ZAR1 system. New data from this study depicted in orange [6–10,17,18,21,32]

**Supplemental Table S1:** Results from Hyphy Contrast-FEL model comparing selection pressure among ZRKs

| RKS1 residue | Alignment Codon | alpha | beta (background) | beta (Foreground) | P-value (overall) | Q-value (overall) |
|---|---|---|---|---|---|---|
| 18 | 1 | 3.127 | 0.658 | 1.826 | 0.692 | 1 |
| 19 | 2 | 2.234 | 0.66 | 89.126 | 0.126 | 1 |
| 20 | 3 | 5.9 | 1.375 | 14.483 | 0.298 | 1 |
| 21 | 4 | 2.656 | 0.661 | 0 | 0.482 | 1 |
| 22 | 5 | 0.629 | 1.09 | 32.273 | 0.058 | 1 |
| 23 | 6 | 3.404 | 0.935 | 192.141 | 0.076 | 1 |
| 24 | 7 | 1.707 | 1.314 | 13.608 | 0.277 | 1 |
| 25 | 8 | 0.488 | 0.227 | 426.372 | 0.002 | 0.084 |
| 26 | 9 | 1.331 | 0.117 | 554.59 | 0 | 0.032 |
| 27 | 10 | 1.277 | 0 | 71.363 | 0.002 | 0.075 |
| 28 | 11 | 0.851 | 0.462 | 391.846 | 0.002 | 0.089 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 29 | 12 | 1.585 | 0.494 | 528.178 | 0.012 | 0.318 |
| 30 | 13 | 3.674 | 0.166 | 0 | 0.759 | 1 |
| 31 | 14 | 2.629 | 0.094 | 81.62 | 0.001 | 0.076 |
| 32 | 15 | 0.79 | 0.28 | 59.669 | 0.003 | 0.091 |
| 33 | 16 | 1.185 | 0.127 | 9.663 | 0.002 | 0.093 |
| 34 | 17 | 0.614 | 0.071 | 0 | 0.815 | 1 |
| 35 | 18 | 0.871 | 0.066 | 0 | 0.818 | 1 |
| 36 | 19 | 1.098 | 0.198 | 0 | 0.677 | 1 |
| 37 | 20 | 1.394 | 0.652 | 0 | 0.58 | 1 |
| 38 | 21 | 1.151 | 0.127 | 0 | 0.629 | 1 |
| 39 | 22 | 2.652 | 0.251 | 0 | 0.663 | 1 |
| 40 | 23 | 0.468 | 0 | 0 | 0.873 | 1 |
| 41 | 24 | 0 | 0.065 | 0 | 0.83 | 1 |
| 42 | 25 | 0.139 | 0.311 | 0 | 0.657 | 1 |
| 43 | 26 | 0 | 0.059 | 0 | 0.817 | 1 |
| 44 | 27 | 0.921 | 0 | 0 | 0.913 | 1 |
| 45 | 28 | 0.572 | 0 | 0 | 0.949 | 0.998 |
| 46 | 29 | 2.045 | 0.066 | 0 | 0.796 | 1 |
| 47 | 30 | 0.502 | 0.649 | 0 | 0.481 | 1 |
| 48 | 31 | 1.51 | 0 | 0 | 0.909 | 1 |
| 49 | 32 | 0.55 | 0.067 | 0 | 0.817 | 1 |
| 50 | 33 | 0.572 | 0.209 | 0 | 0.686 | 1 |
| 51 | 34 | 1.533 | 0.235 | 3.363 | 0.091 | 1 |
| 52 | 35 | 1.36 | 0.131 | 0 | 0.743 | 1 |
| 53 | 36 | 0.55 | 0 | 0 | 0.942 | 1 |
| 54 | 37 | 1.22 | 0.314 | 0 | 0.628 | 1 |
| 55 | 38 | 0 | 0.211 | 3.307 | 0.153 | 1 |
| 56 | 39 | 0.841 | 0 | 0 | 0.934 | 1 |
| 57 | 40 | 0.433 | 0 | 0 | 0.847 | 0.996 |
| 58 | 41 | 2.704 | 0.307 | 0 | 0.593 | 1 |
| 59 | 42 | 1.878 | 0.243 | 0 | 0.654 | 1 |
| 60 | 43 | 0.279 | 0 | 0 | 0.984 | 1 |
| 61 | 44 | 0.209 | 0.477 | 0 | 0.5 | 1 |
| 62 | 45 | 4.86 | 1.315 | 0 | 0.352 | 1 |
| 63 | 46 | 1.487 | 0.264 | 0.749 | 0.83 | 1 |
| 64 | 47 | 2.194 | 0.293 | 0 | 0.681 | 1 |
| 65 | 48 | 0.705 | 0.534 | 5.77 | 0.162 | 1 |
| 66 | 49 | 1.287 | 0.227 | 0 | 0.674 | 1 |
| 70 | 50 | 1.618 | 0.644 | 0 | 0.442 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 71 | 51 | 3.285 | 1.427 | 0 | 0.311 | 1 |
| 72 | 52 | 0.771 | 0.316 | 0 | 0.599 | 1 |
| 73 | 53 | 0.336 | 1.807 | 0 | 0.351 | 1 |
| 74 | 54 | 3.323 | 0.2 | 0 | 0.693 | 1 |
| 75 | 55 | 1.237 | 0 | 0 | 0.904 | 1 |
| 76 | 56 | 0.515 | 0.32 | 0 | 0.613 | 1 |
| 77 | 57 | 0.814 | 0 | 0 | 0.941 | 1 |
| 78 | 58 | 1.865 | 0.546 | 2.141 | 0.51 | 1 |
| 79 | 59 | 1.507 | 0.206 | 0 | 0.686 | 1 |
| 80 | 60 | 0.371 | 0.282 | 0 | 0.635 | 1 |
| 81 | 61 | 0.799 | 0.257 | 0 | 0.652 | 1 |
| 82 | 62 | 2.783 | 0.314 | 0 | 0.597 | 1 |
| 83 | 63 | 1.552 | 0.577 | 3.433 | 0.202 | 1 |
| 84 | 64 | 1.802 | 0.427 | 0 | 0.584 | 1 |
| 85 | 65 | 2.288 | 0.24 | 0 | 0.703 | 1 |
| 86 | 66 | 0 | 0.135 | 0 | 0.742 | 1 |
| 87 | 67 | 0.268 | 0.195 | 0 | 0.686 | 1 |
| 88 | 68 | 0.29 | 0.203 | 56.386 | 0.001 | 0.097 |
| 90 | 69 | 10000 | 2.047 | 0 | 0.768 | 1 |
| 91 | 70 | 2.943 | 1.005 | 0 | 0.633 | 1 |
| 92 | 71 | 10000 | 1.145 | 79.163 | 0.059 | 1 |
| 93 | 72 | 10000 | 2.741 | 0.178 | 0.66 | 1 |
| 94 | 73 | 1.649 | 1.297 | 2.773 | 0.763 | 1 |
| 100 | 74 | 0.714 | 1.29 | 0 | 0.335 | 1 |
| 101 | 75 | 0.806 | 1.354 | 0 | 0.584 | 1 |
| 102 | 76 | 0 | 0.909 | 0 | 0.385 | 1 |
| 103 | 77 | 2.346 | 0.691 | 0 | 0.441 | 1 |
| 104 | 78 | 0.479 | 0.569 | 0 | 0.512 | 1 |
| 105 | 79 | 1.683 | 0.138 | 0 | 0.764 | 1 |
| 106 | 80 | 1.288 | 0.147 | 0 | 0.711 | 1 |
| 107 | 81 | 1.151 | 0.107 | 0 | 0.748 | 1 |
| 108 | 82 | 1.008 | 0 | 0 | 0.915 | 1 |
| 109 | 83 | 1.331 | 0.126 | 0 | 0.725 | 1 |
| 110 | 84 | 0.853 | 0.244 | 0 | 0.667 | 1 |
| 111 | 85 | 0.522 | 0.087 | 0 | 0.828 | 1 |
| 112 | 86 | 0.97 | 0.077 | 0 | 0.822 | 1 |
| 113 | 87 | 0.686 | 0.272 | 0 | 0.627 | 1 |
| 114 | 88 | 1.269 | 0.058 | 0 | 0.794 | 1 |
| 115 | 89 | 20 | 0 | 0 | 0.991 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 116 | 90 | 1.061 | 0.179 | 0 | 0.656 | 1 |
| 117 | 91 | 1.198 | 0 | 0 | 0.904 | 1 |
| 118 | 92 | 67.428 | 0.127 | 0 | 0.747 | 1 |
| 119 | 93 | 0.689 | 0 | 0 | 1 | 1 |
| 120 | 94 | 0.575 | 0.202 | 0 | 0.702 | 1 |
| 121 | 95 | 1.612 | 0.144 | 0 | 0.741 | 1 |
| 122 | 96 | 1.677 | 0.131 | 0 | 0.735 | 1 |
| 123 | 97 | 1.234 | 0 | 0 | 0.872 | 1 |
| 124 | 98 | 0.999 | 0.437 | 0 | 0.598 | 1 |
| 125 | 99 | 0.527 | 0 | 0 | 0.921 | 1 |
| 126 | 100 | 1.242 | 0.114 | 0 | 0.744 | 1 |
| 127 | 101 | 0 | 0 | 0 | 1 | 1 |
| 128 | 102 | 0.97 | 0.074 | 0 | 0.832 | 0.999 |
| 129 | 103 | 0.251 | 0 | 0 | 0.951 | 0.996 |
| 130 | 104 | 3.485 | 0.131 | 0 | 0.738 | 1 |
| 131 | 105 | 0 | 1.244 | 0 | 0.311 | 1 |
| 132 | 106 | 0 | 0.868 | 0 | 0.39 | 1 |
| 133 | 107 | 0.681 | 0 | 0 | 1 | 1 |
| 134 | 108 | 0.305 | 0.072 | 0 | 0.819 | 1 |
| 135 | 109 | 1.707 | 0.161 | 0 | 0.727 | 1 |
| 136 | 110 | 1.111 | 0 | 0 | 0.9 | 1 |
| 137 | 111 | 0.49 | 0.202 | 0 | 0.665 | 1 |
| 142 | 112 | 0.509 | 0.463 | 42.809 | 0.022 | 0.486 |
| 143 | 113 | 0 | 0.629 | 0 | 0.492 | 1 |
| 144 | 114 | 0.595 | 0.64 | 0 | 0.483 | 1 |
| 156 | 115 | 2.04 | 1.222 | 2.741 | 0.871 | 1 |
| 157 | 116 | 5.894 | 1.138 | 0 | 0.403 | 1 |
| 158 | 117 | 7.995 | 1.021 | 5.456 | 0.496 | 1 |
| 159 | 118 | 1.239 | 1.57 | 1.805 | 0.945 | 1 |
| 161 | 119 | 1.555 | 0.057 | 3.497 | 0.022 | 0.518 |
| 162 | 120 | 0 | 0.281 | 0 | 0.64 | 1 |
| 163 | 121 | 0.464 | 0.441 | 0 | 0.543 | 1 |
| 164 | 122 | 1.448 | 0.853 | 0 | 0.458 | 1 |
| 165 | 123 | 1.568 | 0 | 0 | 0.871 | 1 |
| 166 | 124 | 0.438 | 0.259 | 0 | 0.681 | 1 |
| 167 | 125 | 0.563 | 0.062 | 0 | 0.853 | 0.998 |
| 168 | 126 | 0.992 | 0 | 0 | 0.915 | 1 |
| 169 | 127 | 1.663 | 0.124 | 0 | 0.728 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 170 | 128 | 3.635 | 0.131 | 0 | 0.74 | 1 |
| 171 | 129 | 0 | 0.125 | 0 | 0.746 | 1 |
| 172 | 130 | 0.94 | 0.066 | 0 | 0.815 | 1 |
| 173 | 131 | 2.606 | 0 | 0 | 0.974 | 1 |
| 174 | 132 | 0 | 0.259 | 0 | 0.657 | 1 |
| 175 | 133 | 0.981 | 0.06 | 0 | 0.818 | 1 |
| 176 | 134 | 0.549 | 0.271 | 0 | 0.661 | 1 |
| 177 | 135 | 0.162 | 0.278 | 2.527 | 0.254 | 1 |
| 178 | 136 | 1.017 | 0 | 0 | 0.933 | 1 |
| 179 | 137 | 0.162 | 0 | 0 | 0.931 | 1 |
| 180 | 138 | 0.304 | 0 | 0 | 1 | 1 |
| 181 | 139 | 1.567 | 0.123 | 0 | 0.748 | 1 |
| 182 | 140 | 1.291 | 0.06 | 0 | 0.839 | 0.998 |
| 183 | 141 | 2.333 | 0.06 | 0 | 0.808 | 1 |
| 184 | 142 | 3.478 | 0.12 | 0 | 0.739 | 1 |
| 185 | 143 | 1.253 | 0.13 | 0 | 0.717 | 1 |
| 186 | 144 | 0.614 | 0.436 | 0 | 0.555 | 1 |
| 187 | 145 | 0.89 | 0.371 | 0 | 0.575 | 1 |
| 188 | 146 | 0.578 | 0.141 | 0 | 0.728 | 1 |
| 189 | 147 | 0.465 | 0.184 | 0 | 0.661 | 1 |
| 190 | 148 | 0.351 | 0.245 | 0 | 0.619 | 1 |
| 191 | 149 | 0.379 | 0.515 | 0 | 0.492 | 1 |
| 192 | 150 | 0.985 | 0.537 | 0 | 0.609 | 1 |
| 193 | 151 | 1.144 | 0.725 | 0 | 0.454 | 1 |
| 194 | 152 | 1.383 | 0.195 | 0 | 0.73 | 1 |
| 195 | 153 | 0.34 | 3.301 | 0 | 0.452 | 1 |
| 196 | 154 | 1.784 | 0.244 | 0 | 0.637 | 1 |
| 197 | 155 | 0.796 | 0.066 | 0 | 0.816 | 1 |
| 198 | 156 | 0 | 0.064 | 0 | 0.806 | 1 |
| 199 | 157 | 0.642 | 0.268 | 0 | 0.69 | 1 |
| 200 | 158 | 0.89 | 0 | 0 | 0.948 | 1 |
| 201 | 159 | 1.082 | 0.492 | 0 | 0.525 | 1 |
| 202 | 160 | 0 | 0.118 | 0 | 0.745 | 1 |
| 203 | 161 | 0.284 | 0.128 | 0 | 0.751 | 1 |
| 204 | 162 | 0.868 | 0.363 | 0 | 0.623 | 1 |
| 205 | 163 | 1.587 | 0.059 | 0 | 0.803 | 1 |
| 206 | 164 | 0 | 0 | 0 | 1 | 1 |
| 207 | 165 | 0.891 | 0 | 0 | 0.913 | 1 |
| 208 | 166 | 0 | 0.628 | 0 | 0.487 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 209 | 167 | 0.761 | 0.257 | 0 | 0.652 | 1 |
| 210 | 168 | 0.293 | 0.123 | 0 | 0.736 | 1 |
| 211 | 169 | 1.64 | 0.265 | 0 | 0.585 | 1 |
| 212 | 170 | 0.219 | 0.832 | 0 | 0.452 | 1 |
| 213 | 171 | 5.897 | 0.184 | 0 | 0.645 | 1 |
| 214 | 172 | 0.422 | 0.137 | 0 | 0.727 | 1 |
| 215 | 173 | 1.167 | 0.698 | 254.83 | 0.235 | 1 |
| 216 | 174 | 2.896 | 0.139 | 0 | 0.765 | 1 |
| 217 | 175 | 0.34 | 0.116 | 0 | 0.737 | 1 |
| 218 | 176 | 0.313 | 0.064 | 0 | 0.815 | 1 |
| 219 | 177 | 0.3 | 0 | 0 | 1 | 1 |
| 220 | 178 | 1.659 | 0.064 | 0 | 0.809 | 1 |
| 221 | 179 | 0.535 | 0.243 | 347.316 | 0.012 | 0.342 |
| 222 | 180 | 34.161 | 0.48 | 0 | 0.561 | 1 |
| 223 | 181 | 0.923 | 0.071 | 0 | 0.81 | 1 |
| 224 | 182 | 0 | 0.693 | 296.929 | 0.097 | 1 |
| 225 | 183 | 0.527 | 0.285 | 0 | 0.605 | 1 |
| 227 | 184 | 10000 | 1.476 | 0.626 | 0.797 | 1 |
| 228 | 185 | 0.76 | 0.411 | 0 | 0.588 | 1 |
| 229 | 186 | 2.199 | 1.523 | 0 | 0.296 | 1 |
| 230 | 187 | 0.718 | 0.061 | 0 | 0.844 | 1 |
| 231 | 188 | 0.07 | 0.938 | 0 | 0.399 | 1 |
| 232 | 189 | 1.8 | 1.314 | 21.324 | 0.091 | 1 |
| 235 | 190 | 1.649 | 0.724 | 103.129 | 0.033 | 0.671 |
| 236 | 191 | 0 | 0.308 | 0 | 0.632 | 1 |
| 237 | 192 | 1.444 | 0.327 | 0 | 0.586 | 1 |
| 238 | 193 | 1.321 | 0.392 | 5.2 | 0.136 | 1 |
| 239 | 194 | 4.291 | 0 | 0 | 0.895 | 1 |
| 240 | 195 | 0.92 | 1.842 | 66.16 | 0.222 | 1 |
| 241 | 196 | 2.401 | 880.623 | 0.339 | 1 | 1 |
| 242 | 197 | 1.076 | 0.201 | 0 | 0.682 | 1 |
| 243 | 198 | 10000 | 0.378 | 0 | 0.561 | 1 |
| 244 | 199 | 1.263 | 1.435 | 2.276 | 0.823 | 1 |
| 245 | 200 | 0.407 | 0.073 | 0 | 0.814 | 1 |
| 246 | 201 | 2.679 | 0.142 | 0 | 0.73 | 1 |
| 247 | 202 | 1.018 | 0.2 | 0 | 0.683 | 1 |
| 248 | 203 | 0 | 0.531 | 0 | 0.495 | 1 |
| 249 | 204 | 4.146 | 0.367 | 0 | 0.63 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 250 | 205 | 1.549 | 0 | 0 | 0.941 | 1 |
| 251 | 206 | 1.235 | 0.349 | 0 | 0.625 | 1 |
| 252 | 207 | 1.038 | 0.152 | 0 | 0.752 | 1 |
| 253 | 208 | 0.492 | 0.103 | 0 | 0.74 | 1 |
| 254 | 209 | 0 | 0.126 | 0 | 0.778 | 1 |
| 255 | 210 | 0.525 | 0 | 0 | 0.924 | 1 |
| 256 | 211 | 1.282 | 0.55 | 3.976 | 0.267 | 1 |
| 257 | 212 | 4.183 | 0.49 | 0 | 0.544 | 1 |
| 258 | 213 | 0 | 0.103 | 0 | 0.733 | 1 |
| 259 | 214 | 0.634 | 0.517 | 0 | 0.584 | 1 |
| 260 | 215 | 0.35 | 1.226 | 4.165 | 0.518 | 1 |
| 261 | 216 | 0.266 | 0.428 | 0 | 0.607 | 1 |
| 262 | 217 | 1.314 | 0.575 | 0 | 0.546 | 1 |
| 263 | 218 | 1.509 | 0.764 | 0 | 0.439 | 1 |
| 264 | 219 | 0.548 | 0.283 | 0 | 0.638 | 1 |
| 265 | 220 | 1.239 | 0.91 | 0 | 0.427 | 1 |
| 266 | 221 | 1.327 | 0.907 | 8.063 | 0.239 | 1 |
| 267 | 222 | 2.577 | 1.009 | 0 | 0.744 | 1 |
| 268 | 223 | 1.407 | 1.339 | 0 | 0.442 | 1 |
| 274 | 224 | 2.686 | 1.63 | 17.041 | 0.244 | 1 |
| 275 | 225 | 15.635 | 1.547 | 4.455 | 0.587 | 1 |
| 276 | 226 | 1.651 | 5.744 | 64.957 | 0.344 | 1 |
| 277 | 227 | 0 | 5.649 | 2.311 | 0.838 | 1 |
| 278 | 228 | 2.026 | 2.634 | 0 | 0.715 | 1 |
| 279 | 229 | 0.273 | 0.821 | 0 | 0.501 | 1 |
| 280 | 230 | 0.863 | 1.106 | 0 | 0.327 | 1 |
| 281 | 231 | 4.192 | 0.83 | 0 | 0.674 | 1 |
| 284 | 232 | 1.44 | 0.974 | 0 | 0.765 | 1 |
| 285 | 233 | 1.697 | 0.494 | 0 | 0.55 | 1 |
| 286 | 234 | 0 | 0.849 | 4.111 | 0.29 | 1 |
| 287 | 235 | 0 | 1.678 | 9.657 | 0.25 | 1 |
| 288 | 236 | 0.495 | 0.509 | 2.944 | 0.246 | 1 |
| 289 | 237 | 1.158 | 0.494 | 0 | 0.62 | 1 |
| 290 | 238 | 1.833 | 0.592 | 0 | 0.495 | 1 |
| 291 | 239 | 0.166 | 1.591 | 8.021 | 0.395 | 1 |
| 292 | 240 | 2.721 | 0.51 | 0 | 0.69 | 1 |
| 295 | 241 | 1.656 | 0.417 | 0 | 0.619 | 1 |
| 296 | 242 | 0.656 | 0.376 | 0 | 0.569 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 297 | 243 | 0.51 | 0.055 | 0 | 0.81 | 1 |
| 298 | 244 | 0.54 | 0.494 | 0 | 0.501 | 1 |
| 299 | 245 | 0.764 | 1.012 | 0 | 0.511 | 1 |
| 300 | 246 | 0 | 0.247 | 2.143 | 0.217 | 1 |
| 301 | 247 | 1.996 | 0.809 | 0 | 0.53 | 1 |
| 302 | 248 | 1.066 | 0.439 | 0 | 0.691 | 1 |
| 303 | 249 | 137.15 | 0.852 | 0 | 0.37 | 1 |
| 308 | 250 | 0.848 | 0.604 | 0 | 0.573 | 1 |
| 309 | 251 | 3.157 | 0.802 | 3.337 | 0.454 | 1 |
| 310 | 252 | 0.573 | 0.676 | 0 | 0.539 | 1 |
| 311 | 253 | 0.547 | 0.289 | 0 | 0.629 | 1 |
| 312 | 254 | 0.363 | 0.18 | 0 | 0.653 | 1 |
| 313 | 255 | 0 | 0.208 | 0 | 0.684 | 1 |
| 314 | 256 | 1.378 | 0.059 | 0 | 0.818 | 1 |
| 315 | 257 | 2.217 | 0.124 | 0 | 0.747 | 1 |
| 316 | 258 | 0.927 | 0.426 | 3.175 | 0.324 | 1 |
| 317 | 259 | 1.062 | 0.967 | 50.609 | 0.13 | 1 |
| 318 | 260 | 1.935 | 0 | 0 | 0.886 | 1 |
| 319 | 261 | 2.656 | 0.07 | 0 | 0.803 | 1 |
| 320 | 262 | 0.819 | 0.271 | 0 | 0.718 | 1 |
| 321 | 263 | 0 | 0 | 0 | 1 | 1 |
| 322 | 264 | 1.277 | 0 | 0 | 0.901 | 1 |
| 323 | 265 | 10000 | 0.614 | 0 | 0.535 | 1 |
| 324 | 266 | 0.846 | 0.265 | 2.461 | 0.247 | 1 |
| 325 | 267 | 0.635 | 0.922 | 0 | 0.39 | 1 |
| 326 | 268 | 1.513 | 0.39 | 0 | 0.556 | 1 |
| 328 | 269 | 3.148 | 0.362 | 0 | 0.592 | 1 |
| 329 | 270 | 0.517 | 0.671 | 4.24 | 0.279 | 1 |
| 330 | 271 | 0.318 | 0.238 | 9.869 | 0.001 | 0.075 |
| 331 | 272 | 0.674 | 0.329 | 2.78 | 0.208 | 1 |
| 332 | 273 | 3.362 | 0.545 | 3.452 | 0.3 | 1 |
| 333 | 274 | 0 | 0 | 0 | 1 | 1 |
| 334 | 275 | 1.567 | 0.217 | 0 | 0.66 | 1 |
| 335 | 276 | 33.323 | 0.202 | 0 | 0.662 | 1 |
| 336 | 277 | 1.013 | 0.072 | 0 | 0.814 | 1 |
| 337 | 278 | 0.875 | 0.059 | 0 | 0.818 | 1 |
| 338 | 279 | 0.324 | 0 | 0 | 0.982 | 1 |
| 339 | 280 | 1.69 | 0 | 0 | 0.909 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 340 | 281 | 0.388 | 0 | 0 | 0.891 | 1 |
| 341 | 282 | 1.095 | 0.128 | 0 | 0.742 | 1 |
| 342 | 283 | 0.087 | 1.155 | 3.324 | 0.845 | 0.997 |
| 343 | 284 | 1.11 | 0.066 | 0 | 0.812 | 1 |
| 344 | 285 | 8.347 | 0.137 | 4.197 | 0.045 | 0.867 |
| 345 | 286 | 0 | 0.574 | 0 | 0.508 | 1 |
| 346 | 287 | 0.536 | 1.451 | 2.381 | 0.747 | 1 |
| 347 | 288 | 0.703 | 1.262 | 62.609 | 0.335 | 1 |

*alpha: synonymous substitution rate, beta: non-synonymous substitution rate. Foreground selection is gZRK7, while background is all other ZRKs. Result considered significant if Q-value <0.2

**Supplemental Table S2:** Results from Hyphy Contrast-FEL model comparing ZRKs with WAKLs and PBLs

| RKS1 residue | Alignment Codon | alpha | beta (Foreground) | beta (background) | P-value (overall) | Q-value (overall) |
|---|---|---|---|---|---|---|
| 14 | 1 | 3.299 | 0.817 | 0.339 | 0.243 | 0.48 |
| 15 | 2 | 1.803 | 0.755 | 0.768 | 0.912 | 0.967 |
| 16 | 3 | 0.612 | 0.739 | 0.742 | 0.947 | 0.973 |
| 17 | 4 | 1.147 | 1.131 | 0.502 | 0.12 | 0.278 |
| 18 | 5 | 0.92 | 0.715 | 0.849 | 0.783 | 0.907 |
| 19 | 6 | 0.883 | 0.583 | 0.351 | 0.488 | 0.697 |
| 20 | 7 | 1.406 | 1.405 | 0.985 | 0.601 | 0.798 |
| 24 | 8 | 5.485 | 2.285 | 0.332 | 0.017 | 0.062 |
| 25 | 9 | 3.04 | 0.571 | 0.958 | 0.405 | 0.654 |
| 26 | 10 | 2.344 | 0.193 | 0.59 | 0.094 | 0.233 |
| 27 | 11 | 1.255 | 0.323 | 0.512 | 0.474 | 0.689 |
| 28 | 12 | 1.807 | 0.326 | 0.439 | 0.739 | 0.891 |
| 29 | 13 | 1.613 | 0.548 | 0.603 | 0.841 | 0.933 |
| 30 | 14 | 2.3 | 0.626 | 0.947 | 0.512 | 0.724 |
| 34 | 15 | 2.933 | 0.331 | 0.514 | 0.565 | 0.774 |
| 35 | 16 | 1.74 | 0.102 | 0.365 | 0.135 | 0.302 |
| 36 | 17 | 2.764 | 0.186 | 0.401 | 0.316 | 0.569 |
| 37 | 18 | 0.871 | 0.701 | 0.709 | 0.951 | 0.973 |
| 38 | 19 | 1.266 | 0 | 0.486 | 0.017 | 0.062 |
| 39 | 20 | 1.441 | 0.147 | 0.461 | 0.219 | 0.45 |
| 40 | 21 | 0.567 | 0 | 0.672 | 0.001 | 0.005 |
| 41 | 22 | 1.513 | 0.098 | 0.359 | 0.129 | 0.294 |
| 42 | 23 | 0.442 | 0.348 | 0.512 | 0.593 | 0.791 |
| 46 | 24 | 1.383 | 0 | 0.189 | 0.059 | 0.161 |
| 47 | 25 | 1.099 | 0.609 | 0.267 | 0.251 | 0.489 |
| 48 | 26 | 8.873 | 0 | 0.008 | 0.675 | 0.853 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 49 | 27 | 0.788 | 0.105 | 0.104 | 0.973 | 0.976 |
| 50 | 28 | 0.638 | 0.225 | 0.143 | 0.577 | 0.782 |
| 51 | 29 | 1.393 | 0.38 | 0.346 | 0.876 | 0.952 |
| 52 | 30 | 0.671 | 0.1 | 0.053 | 0.592 | 0.795 |
| 53 | 31 | 0.584 | 0 | 0.044 | 0.352 | 0.606 |
| 54 | 32 | 1.223 | 0.524 | 0.133 | 0.071 | 0.182 |
| 55 | 33 | 0.742 | 0.281 | 0.251 | 0.566 | 0.771 |
| 56 | 34 | 0.705 | 0 | 0.052 | 0.308 | 0.562 |
| 57 | 35 | 1.041 | 0 | 0 | 0.849 | 0.93 |
| 58 | 36 | 0.974 | 0.425 | 0.191 | 0.255 | 0.494 |
| 59 | 37 | 1.247 | 0.37 | 0.111 | 0.068 | 0.179 |
| 60 | 38 | 0.615 | 0 | 0.008 | 0.68 | 0.856 |
| 61 | 39 | 0.929 | 0.362 | 0.205 | 0.497 | 0.706 |
| 62 | 40 | 1.595 | 1.245 | 0.375 | 0.053 | 0.158 |
| 63 | 41 | 1.618 | 0.087 | 0.28 | 0.189 | 0.399 |
| 64 | 42 | 0.803 | 0.136 | 0.191 | 0.744 | 0.893 |
| 65 | 43 | 1.026 | 0.568 | 0.219 | 0.16 | 0.35 |
| 66 | 44 | 0.948 | 0.135 | 0.125 | 0.939 | 0.98 |
| 67 | 45 | 1.296 | 1.231 | 0.019 | 0 | 0 |
| 68 | 46 | 0.244 | 2.376 | 0.044 | 0 | 0 |
| 73 | 47 | 2.494 | 0.578 | 0.026 | 0 | 0.001 |
| 74 | 48 | 1.027 | 0.598 | 0.021 | 0 | 0.001 |
| 75 | 49 | 1.786 | 0 | 0.072 | 0.285 | 0.531 |
| 76 | 50 | 0.599 | 0.218 | 0.034 | 0.084 | 0.212 |
| 77 | 51 | 1.177 | 0 | 0 | 0.846 | 0.931 |
| 78 | 52 | 1.253 | 0.612 | 0.207 | 0.104 | 0.249 |
| 79 | 53 | 1.147 | 0 | 0.154 | 0.139 | 0.307 |
| 80 | 54 | 0.461 | 0.412 | 0.159 | 0.167 | 0.364 |
| 81 | 55 | 1.085 | 0.289 | 0.212 | 0.655 | 0.845 |
| 82 | 56 | 1.188 | 0.199 | 0.318 | 0.665 | 0.845 |
| 83 | 57 | 1.178 | 0.237 | 0.188 | 0.768 | 0.91 |
| 84 | 58 | 0.343 | 0.103 | 0.078 | 0.829 | 0.936 |
| 85 | 59 | 2.433 | 0.114 | 0.025 | 0.265 | 0.505 |
| 86 | 60 | 0.462 | 0.099 | 0.04 | 0.454 | 0.675 |
| 87 | 61 | 0.424 | 0.198 | 0 | 0.005 | 0.021 |
| 88 | 62 | 0.89 | 0.087 | 0.145 | 0.622 | 0.818 |
| 102 | 63 | 0.483 | 143.419 | 0.249 | 0 | 0 |
| 103 | 64 | 0.815 | 1.218 | 0.045 | 0.001 | 0.005 |
| 104 | 65 | 1.793 | 1.139 | 0.03 | 0 | 0.002 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 105 | 66 | 0.303 | 0.111 | 0.147 | 0.763 | 0.908 |
| 106 | 67 | 1.832 | 0.094 | 0.194 | 0.439 | 0.671 |
| 107 | 68 | 0.963 | 0.172 | 0.008 | 0.015 | 0.059 |
| 108 | 69 | 0.956 | 0 | 0.074 | 0.225 | 0.455 |
| 109 | 70 | 1.23 | 0.092 | 0.249 | 0.306 | 0.563 |
| 110 | 71 | 0.618 | 0.228 | 0.234 | 0.96 | 0.978 |
| 111 | 72 | 1.551 | 0.135 | 0.011 | 0.108 | 0.253 |
| 112 | 73 | 0.833 | 0 | 0.058 | 0.332 | 0.582 |
| 113 | 74 | 2.069 | 0.333 | 0.144 | 0.26 | 0.499 |
| 114 | 75 | 1.784 | 0 | 0.131 | 0.101 | 0.246 |
| 115 | 76 | 1.036 | 0 | 0.218 | 0.021 | 0.072 |
| 117 | 77 | 3.447 | 0 | 0.007 | 0.682 | 0.854 |
| 118 | 78 | 1.567 | 0.094 | 0.084 | 0.919 | 0.967 |
| 119 | 79 | 0.819 | 0 | 0.029 | 0.4 | 0.657 |
| 120 | 80 | 0.83 | 0.136 | 0.04 | 0.36 | 0.607 |
| 121 | 81 | 0.726 | 0.114 | 0.021 | 0.229 | 0.461 |
| 122 | 82 | 1.905 | 0.197 | 0.09 | 0.368 | 0.616 |
| 123 | 83 | 0.636 | 0 | 0.031 | 0.48 | 0.69 |
| 124 | 84 | 0.732 | 0.477 | 0.108 | 0.063 | 0.17 |
| 125 | 85 | 0.87 | 0 | 0 | 0.883 | 0.952 |
| 126 | 86 | 1.36 | 0.196 | 0.039 | 0.102 | 0.246 |
| 127 | 87 | 0.724 | 0 | 0.008 | 0.685 | 0.85 |
| 128 | 88 | 0.967 | 0.115 | 0.159 | 0.754 | 0.901 |
| 129 | 89 | 0.809 | 0 | 0.054 | 0.305 | 0.565 |
| 130 | 90 | 1.138 | 0.1 | 0.137 | 0.768 | 0.906 |
| 131 | 91 | 0.565 | 1.637 | 0.166 | 0 | 0.002 |
| 132 | 92 | 0.854 | 0.934 | 0.152 | 0.008 | 0.035 |
| 133 | 93 | 0.898 | 0 | 0.033 | 0.384 | 0.635 |
| 134 | 94 | 0.436 | 0.119 | 0.116 | 0.947 | 0.977 |
| 135 | 95 | 1.318 | 0.125 | 0.01 | 0.11 | 0.257 |
| 136 | 96 | 0.942 | 0 | 0.01 | 0.682 | 0.85 |
| 137 | 97 | 1.928 | 0.109 | 0.031 | 0.335 | 0.584 |
| 138 | 98 | 0.746 | 0.096 | 0.026 | 0.341 | 0.59 |
| 139 | 99 | 0.847 | 0.72 | 0.17 | 0.051 | 0.155 |
| 140 | 100 | 0.45 | 0.351 | 0.066 | 0.037 | 0.117 |
| 141 | 101 | 1.215 | 0 | 0.167 | 0.075 | 0.192 |
| 142 | 102 | 4.601 | 0.479 | 0.144 | 0.07 | 0.183 |
| 143 | 103 | 0.699 | 0.256 | 0.057 | 0.877 | 0.949 |
| 153 | 104 | 1.345 | 0.366 | 0.215 | 0.539 | 0.754 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 154 | 105 | 1.136 | 2.589 | 0.351 | 0.03 | 0.103 |
| 155 | 106 | 1.417 | 1.602 | 0.505 | 0.048 | 0.146 |
| 159 | 107 | 13.192 | 0.881 | 0.193 | 0.034 | 0.109 |
| 160 | 108 | 1.481 | 0.646 | 0.074 | 0.004 | 0.02 |
| 161 | 109 | 0.938 | 0.292 | 0.146 | 0.323 | 0.574 |
| 162 | 110 | 0 | 0.479 | 0.02 | 0.001 | 0.004 |
| 163 | 111 | 0.766 | 0.348 | 0.481 | 0.655 | 0.848 |
| 164 | 112 | 0.708 | 1.178 | 0.173 | 0.006 | 0.027 |
| 165 | 113 | 0.983 | 0 | 0 | 0.83 | 0.933 |
| 166 | 114 | 0.881 | 0.231 | 0.091 | 0.479 | 0.692 |
| 167 | 115 | 0.628 | 0.097 | 0.071 | 0.771 | 0.906 |
| 168 | 116 | 1.093 | 0 | 0.017 | 0.553 | 0.77 |
| 169 | 117 | 1.714 | 0.101 | 0.042 | 0.472 | 0.69 |
| 170 | 118 | 1.042 | 0.084 | 0.27 | 0.245 | 0.482 |
| 171 | 119 | 0.586 | 0.103 | 0.094 | 0.891 | 0.957 |
| 172 | 120 | 1.744 | 0.102 | 0.055 | 0.604 | 0.798 |
| 173 | 121 | 1.442 | 0 | 0 | 0.84 | 0.936 |
| 174 | 122 | 0.402 | 0.269 | 0.15 | 0.447 | 0.676 |
| 175 | 123 | 1.377 | 0.098 | 0.027 | 0.317 | 0.566 |
| 176 | 124 | 0.75 | 0.306 | 0.039 | 0.048 | 0.147 |
| 177 | 125 | 0.682 | 0.231 | 0.209 | 0.91 | 0.969 |
| 178 | 126 | 1.118 | 0 | 0.009 | 0.696 | 0.86 |
| 179 | 127 | 0.759 | 0 | 0.01 | 0.652 | 0.849 |
| 180 | 128 | 0.814 | 0 | 0 | 0.919 | 0.963 |
| 181 | 129 | 1.74 | 0.099 | 0.205 | 0.437 | 0.672 |
| 182 | 130 | 0.836 | 0.1 | 0.215 | 0.415 | 0.653 |
| 187 | 131 | 0.406 | 0.429 | 0.074 | 0.016 | 0.061 |
| 188 | 132 | 0.792 | 0.109 | 0.054 | 0.563 | 0.779 |
| 189 | 133 | 0.639 | 0.188 | 0.026 | 0.056 | 0.159 |
| 190 | 134 | 0.699 | 0.424 | 0 | 0 | 0 |
| 191 | 135 | 0.377 | 0.511 | 0.015 | 0 | 0.001 |
| 192 | 136 | 0.421 | 0.575 | 0.103 | 0.057 | 0.159 |
| 193 | 137 | 1.435 | 0.829 | 0.019 | 0 | 0.001 |
| 194 | 138 | 0.844 | 0.099 | 0.128 | 0.808 | 0.92 |
| 195 | 139 | 1.324 | 4.349 | 0.099 | 0 | 0 |
| 196 | 140 | 2.4 | 0.376 | 0 | 0.003 | 0.014 |
| 197 | 141 | 0.887 | 0 | 0.017 | 0.564 | 0.777 |
| 198 | 142 | 1.005 | 0 | 0.032 | 0.425 | 0.665 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 199 | 143 | 0.98 | 0.41 | 0.011 | 0.004 | 0.021 |
| 200 | 144 | 0.911 | 0 | 0.03 | 0.408 | 0.654 |
| 201 | 145 | 0.423 | 0.597 | 0.291 | 0.331 | 0.584 |
| 202 | 146 | 0.709 | 0.09 | 0.206 | 0.373 | 0.621 |
| 203 | 147 | 0.591 | 0.113 | 0.178 | 0.643 | 0.841 |
| 204 | 148 | 1.41 | 0.35 | 0.145 | 0.224 | 0.457 |
| 205 | 149 | 1.315 | 0.097 | 0.07 | 0.777 | 0.908 |
| 206 | 150 | 0.727 | 0 | 0 | 0.969 | 0.976 |
| 207 | 151 | 0.655 | 0 | 0.04 | 0.413 | 0.659 |
| 208 | 152 | 0.788 | 0.862 | 0.031 | 0.001 | 0.007 |
| 209 | 153 | 1.695 | 0.678 | 0 | 0 | 0.002 |
| 210 | 154 | 1.371 | 0.185 | 0 | 1 | 1 |
| 211 | 155 | 0.928 | 0.203 | 0.028 | 0.055 | 0.161 |
| 212 | 156 | 1.8 | 0.999 | 0.073 | 0 | 0.002 |
| 213 | 157 | 0.815 | 0.103 | 0.047 | 0.52 | 0.731 |
| 214 | 158 | 0.433 | 0.106 | 0.07 | 0.717 | 0.873 |
| 215 | 159 | 2.617 | 1.092 | 0.219 | 0.014 | 0.057 |
| 216 | 160 | 1.391 | 0.107 | 0.079 | 0.794 | 0.916 |
| 217 | 161 | 0.489 | 0.199 | 0.03 | 0.058 | 0.16 |
| 218 | 162 | 1.127 | 0.075 | 0.308 | 0.174 | 0.373 |
| 219 | 163 | 0.428 | 0 | 0.095 | 0.198 | 0.413 |
| 220 | 164 | 0.807 | 0 | 0.23 | 0.036 | 0.116 |
| 221 | 165 | 2.834 | 2.026 | 0.179 | 0.001 | 0.006 |
| 222 | 166 | 0.435 | 42.994 | 0.082 | 0 | 0 |
| 223 | 167 | 1.063 | 0.434 | 0.056 | 0.017 | 0.061 |
| 224 | 168 | 0.621 | 1.418 | 0.178 | 0.001 | 0.004 |
| 225 | 169 | 0.886 | 1.754 | 0.035 | 0 | 0 |
| 228 | 170 | 0.382 | 0.64 | 0.017 | 0.001 | 0.004 |
| 229 | 171 | 0.206 | 11.895 | 0.085 | 0 | 0 |
| 230 | 172 | 0.801 | 0.106 | 0 | 0.024 | 0.082 |
| 231 | 173 | 0.452 | 1.661 | 0 | 0 | 0 |
| 232 | 174 | 0.661 | 1.906 | 0.221 | 0 | 0.002 |
| 235 | 175 | 0.703 | 0.981 | 0.099 | 0.002 | 0.011 |
| 236 | 176 | 0.435 | 0.461 | 0.014 | 0 | 0.002 |
| 237 | 177 | 0.82 | 0.347 | 0.013 | 0.002 | 0.01 |
| 238 | 178 | 0.737 | 0.114 | 0.043 | 0.433 | 0.671 |
| 239 | 179 | 2.596 | 0 | 0 | 0.944 | 0.978 |
| 240 | 180 | 1.683 | 2.118 | 0.211 | 0 | 0.004 |
| 241 | 181 | 2.783 | 3.74 | 0.197 | 0 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 242 | 182 | 0.491 | 0.103 | 0.077 | 0.803 | 0.918 |
| 243 | 183 | 0.486 | 0.966 | 0.041 | 0 | 0 |
| 244 | 184 | 30.274 | 2.116 | 0.129 | 0 | 0.001 |
| 245 | 185 | 1.607 | 0.121 | 0.046 | 0.443 | 0.673 |
| 246 | 186 | 1.079 | 0.106 | 0.042 | 0.433 | 0.674 |
| 247 | 187 | 2.083 | 0.317 | 0.25 | 0.726 | 0.88 |
| 248 | 188 | 1.102 | 0.463 | 0.071 | 0.01 | 0.042 |
| 249 | 189 | 2.056 | 0.632 | 0.038 | 0.004 | 0.021 |
| 250 | 190 | 0.919 | 0 | 0 | 0.915 | 0.966 |
| 251 | 191 | 0.784 | 0.558 | 0.021 | 0 | 0.002 |
| 252 | 192 | 0.938 | 0.121 | 0.042 | 0.402 | 0.657 |
| 253 | 193 | 1.096 | 0 | 0.048 | 0.282 | 0.534 |
| 254 | 194 | 0.695 | 0.205 | 0.033 | 0.068 | 0.181 |
| 255 | 195 | 0.886 | 0 | 0 | 0.896 | 0.958 |
| 256 | 196 | 1.169 | 0.791 | 0.02 | 0 | 0 |
| 257 | 197 | 0.774 | 1.019 | 0.056 | 0 | 0.003 |
| 258 | 198 | 2.182 | 0.081 | 0.011 | 0.183 | 0.388 |
| 259 | 199 | 1.585 | 0.135 | 0.132 | 0.964 | 0.975 |
| 260 | 200 | 0.372 | 2.239 | 0.014 | 0 | 0 |
| 261 | 201 | 0.533 | 0.466 | 0.094 | 0.056 | 0.158 |
| 262 | 202 | 1.521 | 0.575 | 0.082 | 0.031 | 0.104 |
| 263 | 203 | 0.658 | 0.995 | 0.07 | 0 | 0.003 |
| 264 | 204 | 0.419 | 0.504 | 0.009 | 0 | 0.002 |
| 265 | 205 | 1.097 | 1.157 | 0.109 | 0.001 | 0.004 |
| 266 | 206 | 1.778 | 0.924 | 0.108 | 0.001 | 0.005 |
| 267 | 207 | 0.873 | 2.109 | 0.1 | 0 | 0.001 |
| 268 | 208 | 0.735 | 2.167 | 0.304 | 0.005 | 0.025 |
| 269 | 209 | 0.372 | 25.253 | 0.056 | 0 | 0 |
| 270 | 210 | 1.974 | 1.646 | 0.309 | 0.008 | 0.034 |
| 271 | 211 | 0.698 | 3.837 | 0.348 | 0 | 0.002 |
| 272 | 212 | 2.045 | 2.675 | 0.083 | 0 | 0 |
| 273 | 213 | 0.471 | 1.454 | 0.211 | 0.008 | 0.035 |
| 274 | 214 | 2.824 | 1.453 | 0.374 | 0.055 | 0.16 |
| 275 | 215 | 0.773 | 3.253 | 0.38 | 0 | 0.004 |
| 276 | 216 | 1.172 | 18.061 | 0.176 | 0 | 0 |
| 288 | 217 | 0.488 | 0.972 | 0.123 | 0.001 | 0.004 |
| 289 | 218 | 0.905 | 0.979 | 0.238 | 0.045 | 0.141 |
| 290 | 219 | 0.952 | 0.519 | 0.109 | 0.025 | 0.085 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 292 | 220 | 1.78 | 0.244 | 0.187 | 0.778 | 0.906 |
| 293 | 221 | 0.884 | 0.911 | 0.443 | 0.31 | 0.562 |
| 294 | 222 | 0.966 | 0.836 | 0.185 | 0.016 | 0.061 |
| 295 | 223 | 1.292 | 0.866 | 0.099 | 0.002 | 0.01 |
| 296 | 224 | 0.876 | 0.4 | 0.145 | 0.283 | 0.532 |
| 297 | 225 | 0.959 | 0.198 | 0.007 | 0.009 | 0.038 |
| 298 | 226 | 0.852 | 1.022 | 0.15 | 0.001 | 0.005 |
| 302 | 227 | 1.73 | 0.586 | 0.188 | 0.104 | 0.248 |
| 303 | 228 | 2.818 | 1.065 | 0.208 | 0.005 | 0.025 |
| 304 | 229 | 0.414 | 1.346 | 0.143 | 0 | 0.001 |
| 305 | 230 | 1.492 | 0.102 | 0.225 | 0.403 | 0.654 |
| 309 | 231 | 1.342 | 0.979 | 0.386 | 0.134 | 0.302 |
| 310 | 232 | 0.914 | 0.869 | 0.219 | 0.032 | 0.106 |
| 311 | 233 | 0.862 | 0.415 | 0.167 | 0.207 | 0.428 |
| 312 | 234 | 1.328 | 0.179 | 0.142 | 0.8 | 0.919 |
| 313 | 235 | 0.896 | 0.053 | 0.374 | 0.1 | 0.246 |
| 314 | 236 | 0.579 | 0.097 | 0.221 | 0.359 | 0.609 |
| 315 | 237 | 1.08 | 0.092 | 0.187 | 0.471 | 0.695 |
| 316 | 238 | 0.485 | 0.497 | 0.075 | 0.016 | 0.061 |
| 317 | 239 | 0.93 | 1.422 | 0.281 | 0.054 | 0.16 |
| 318 | 240 | 1.98 | 0 | 0.086 | 0.234 | 0.465 |
| 319 | 241 | 1.49 | 0 | 0.052 | 0.353 | 0.603 |
| 320 | 242 | 0.688 | 0.44 | 0.253 | 0.447 | 0.672 |
| 321 | 243 | 0.754 | 0 | 0.109 | 0.132 | 0.3 |
| 322 | 244 | 1.197 | 0 | 0 | 0.843 | 0.932 |
| 323 | 245 | 1.617 | 0.533 | 0.152 | 0.087 | 0.219 |
| 324 | 246 | 1.482 | 0.131 | 0.159 | 0.838 | 0.937 |
| 325 | 247 | 1.322 | 0.335 | 0.461 | 0.7 | 0.86 |
| 328 | 248 | 1.127 | 0.146 | 0.143 | 0.962 | 0.977 |
| 329 | 249 | 1.556 | 0.331 | 0.536 | 0.712 | 0.871 |
| 330 | 250 | 1.56 | 0.124 | 0.017 | 0.17 | 0.367 |
| 331 | 251 | 0.639 | 0.514 | 0 | 0 | 0 |
| 332 | 252 | 1.49 | 0.465 | 0.4 | 0.852 | 0.93 |
| 333 | 253 | 0 | 0 | 0.019 | 0.472 | 0.693 |
| 334 | 254 | 0.967 | 0.135 | 0.281 | 0.448 | 0.669 |
| 335 | 255 | 10000 | 0.101 | 0.15 | 0.661 | 0.848 |
| 336 | 256 | 0.939 | 0.12 | 0.043 | 0.414 | 0.656 |
| 337 | 257 | 0.891 | 0.095 | 0.088 | 0.94 | 0.977 |
| 338 | 258 | 0.688 | 0 | 0.41 | 0.006 | 0.025 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 339 | 259 | 0.826 | 0 | 0.189 | 0.058 | 0.158 |
| 340 | 260 | 1.194 | 0 | 0.01 | 0.663 | 0.847 |
| 341 | 261 | 0.667 | 0.195 | 0.123 | 0.577 | 0.778 |
| 342 | 262 | 0.067 | 1.558 | 0.365 | 0.019 | 0.067 |
| 343 | 263 | 1.16 | 0.099 | 0.115 | 0.811 | 0.919 |

*alpha: synonymous substitution rate, beta: non-synonymous substitution rate. For foreground and background selection, see Supplemental Figure S6. Result considered significant if Q-value <0.2

## CONCLUSION

Determining ZAR1 interaction specificity enabled rational manipulation of interaction strength with ZAR1 and a ZRK kinase. This is a key step towards rational design of ZAR1 recognition pathways for *P. syringae* resistance.

# CHAPTER IV: Molecular dissection of the ZED1 pseudokinase enables expansion of HopZ1a recognition to *Solanum lycopersicum* ZAR1

## PREFACE

This chapter was done in collaboration with several researchers. I performed the co-immunoprecipitations, yeast-two hybrids, co-immunoprecipitations, and structural and docking analyses. I performed transient gene expressions and HR assays with Maël Baudin and Derek Xiang, the confocal microscopy with Derek Xiang and analyzed the data with Maël Baudin. Lung-Yu Liang, Weiwen Dai, James M. Murphy and Isabelle L. Lucet conducted the ZED1 thermal shift assay. I wrote the manuscript with Maël Baudin and Jennifer D. Lewis, and with assistance from Derek Xiang, Lung-Yu Liang, Weiwen Dai, James M. Murphy, Isabelle L. Lucet, and Jana A. Hassan. The work has been reproduced below with the permission of all authors.

This work uses diversity in a ZAR1 interacting protein, ZED1, to identify determinants for ZAR1 recognition of HopZ1a. This worked enables recognition of HopZ1a through tomato ZAR1, a key step towards rational design of ZAR1 recognition pathways for *P. syringae* resistance.


## Abstract

The highly conserved angiosperm immune receptor ZAR1 is a bacterial pathogen recognition hub, which mediates resistance by guarding host kinases for modification by pathogen effectors. The pseudokinase ZED1 is the only known ZAR1-guarded protein that interacts directly with a pathogen effector, HopZ1a from the bacterial pathogen *Pseudomonas syringae*, making it a promising system for rational design of effector recognition for plant immunity. Here, we conduct the most in-depth molecular analysis of ZED1 to date. We generated a library of 164 random ZED1 mutants and identified 50 mutants that were unable to recognize the effector HopZ1a when transiently expressed in *Nicotiana benthamiana*. Based on our random mutants, we generated a library of 27 point mutants and found evidence of minor functional divergence between *A. thaliana* and *N. benthamiana* ZAR1 orthologs. We leveraged our point mutant library to identify regions in ZED1 that are critical for ZAR1 and HopZ1a interactions, and identified two likely ZED1-HopZ1a binding conformations. We explored ZED1 nucleotide and cation binding activity, and showed that ZED1 is a catalytically dead pseudokinase, functioning solely as an allosteric regulator upon effector recognition. We used our library of ZED1 point mutants to identify the ZED1 activation loop regions as the most likely cause of interspecies ZAR1-ZED1 incompatibility. Finally, we identified a mutation that abolished ZAR1-ZED1 interspecies incompatibility while retaining the ability to mediate HopZ1a recognition, which enabled recognition of HopZ1a through *Solanum lycopersicum* ZAR1. This is the first example of expanded effector recognition through a ZAR1 ortholog from a non-model species.


## Introduction

The plant immune system functions in two interconnected tiers [1]. Plants recognize conserved microbe associated molecular patterns through pattern recognition receptors at the cell surface, leading to pattern-triggered immunity (PTI) [2]. Some pathogens have evolved the ability to secrete effector proteins into plant cells to suppress PTI and promote virulence. Plants have

evolved the nucleotide-binding leucine-rich repeat (NLR) family of intracellular immune receptors that recognize bacterial effectors and lead to effector-triggered immunity (ETI). ETI is often associated with a form of programmed cell death termed the hypersensitive response (HR) [3–5].

The *Arabidopsis thaliana* NLR HopZ-Activated Resistance 1 (AtZAR1) is a resistance hub with functionally conserved homologs across the plant kingdom [6–9]. AtZAR1 mediates ETI against 5 of the 19 known *Pseudomonas syringae* effector families recognized by *A. thaliana*, the most of any characterized NLR [10–12]. In addition, AtZAR1 has been shown to confer resistance to *Xanthomonas campestris* pv. *campestris* through indirect recognition of the effector AvrAC [13]. The broad effector recognition repertoire of AtZAR1 is made possible by monitoring kinases for effector modification. AtZAR1 interacts directly with ZED1, and ZED1-related kinases (ZRKs, RLCK XII-2) in the absence of pathogen effectors [11–14]. Most ZRKs require an additional tier of PBS1-like kinases (PBLs), that are modified by pathogen effectors and interact with ZRKs upon effector modification [13,15]. ZED1 is the only kinase that has been shown to interact directly with AtZAR1 and its recognized effector, HopZ1a [12,13,15]. Furthermore, HopZ1a-triggered ETI requires acetylation of ZED1 by HopZ1a [12]. Although no PBLs have been shown to be fully required for ZED1 recognition of HopZ1a, several ZED1-PBL and HopZ1a-PBL interactions have been characterized [16,17]. Perturbed interactions between ZED1 and AtZAR1, or ZED1 and HopZ1a result in a loss of ZED1-mediated ETI [7,12,18].

The ZRK family likely evolved from the pathogen-targeted Wall Associated Kinase (WAK) family via the loss of receptor domains [9]. While some ZRKs, such as ZRK10 have been shown to possess kinase activity, ZED1 appears to be a pseudokinase [12,16], lacking several key residues that confer the catalytic activity of canonical kinases. The HRD motif, which is responsible for phospho-transfer in canonical kinases, has an asparagine substitution in place of aspartic acid in ZED1. The DFG motif, a key regulator of the kinase activation loop and binding of $Mg^{2+}$ for phosphotransfer, has a tryptophan substituted for the glycine in ZED1. The ZED1 activation loop is devoid of phospho-accepting serine or threonine residues, a key feature of canonical Ser/Thr kinase activation loops. The substitution of valine to a phospho-accepting threonine residue in the ZED1 activation loop ($ZED1^{V212T}$) leads to an autoactive HR phenotype [16]. While restoring the HRD motif in ZED1 by substituting asparagine for the canonical aspartic acid ($ZED1^{N173D}$) does not have a phenotypic effect, further mutating the asparagine in the HRD motif to a serine ($ZED1^{N173S)}$ results in an autoactive HR phenotype [16,19]. Finally, restoring the DFG motif does not result in phenotypic effects [16]. While ZED1 does not possess the capacity to phosphorylate substrates [12], it is possible that ZED1 still retains some cation or nucleotide binding capacities that are important for effector recognition and resistance.

The recent cryo-EM structure of AtZAR1 in complex with RKS1 in inactive, intermediate, or active pentameric conformations provides a structural template for understanding ZAR1-ZRK interactions and ZAR1 function. RKS1 is a member of the ZRK family and interacts with AtZAR1 in the absence of pathogen effectors. It mediates ETI in response to the uridylation of PBL2 by the effector AvrAC [13]. The C-terminal AtZAR1 leucine-rich repeat (hereafter $AtZAR1^{LRR}$) is the main region of contact for RKS1 interactions. We and others have previously shown that $AtZAR1^{LRR}$ is the determinant of AtZAR1 interaction specificity, and that disrupted $AtZAR1^{LRR}$-ZRK interactions lead to a loss of ETI [7,13,14,20]. Upon uridylation, PBL2 interacts with RKS1, causing the RKS1 activation loop to undergo conformational changes leading to a clash with the AtZAR1 nucleotide binding domain (hereafter $AtZAR1^{NBD}$), and assembly of AtZAR1 into a wheel-like pentamer termed the

"resistosome" [21,22]. The resistosome forms a cation channel at the plasma membrane via the AtZAR1 coiled-coil domain (hereafter AtZAR1$^{CC}$).

The functional conservation of ZAR1 across angiosperms makes it unique among NLRs. The domain structure of ZAR1 orthologs is highly conserved and ZAR1 orthologs can function in effector recognition with ZRK orthologs within species (intraspecies) or between species (interspecies). We have previously shown that ZED1 from *A. thaliana*, can mediate ETI when transiently co-expressed with HopZ1a in another model species, *Nicotiana benthamiana*, and that this ETI is dependent on the *N. benthamiana* ZAR1 ortholog (hereafter NbZAR1) [7]. ZED1 can mediate ETI through AtZAR1 and NbZAR1 despite having only 58% similarity between the two orthologs at the protein level [7]. NbZAR1 can also recognize *Xanthomonas perforans* through indirect recognition of XopJ4 and HopZ5 mediated by *N. benthamiana* ZRKs [8,23]. Across angiosperms, every known species with a ZAR1 ortholog also possesses ZRKs, further supporting that the mechanism of ZAR1 effector recognition through ZRKs may be conserved outside of the Brassicaceae and Solanaceae [9,24]. While some interspecies ZAR1-ZRK pairings constitute a functional ETI system, other interspecies pairings may result in an autoactive HR phenotype in the absence of effectors. We have shown that transiently expressing *A. thaliana* RKS1 or ZRK13 in *N. benthamiana* results in a NbZAR1-dependent HR phenotype in the absence of effectors [14]. In addition, transiently expressing at least six diverse ZAR1 orthologs in *A. thaliana zar1* protoplasts results in cell death in the absence of effectors [9]. While this autoactive HR phenotype is dependent on ZRKs, the exact mechanism that causes activation is not understood.

Understanding the molecular mechanisms of HopZ1a recognition through ZED1 may allow for simplified rational design of effector recognition through ZAR1, without a second tier of PBLs. In addition, understanding the molecular mechanisms that lead to autoactive HR phenotypes resulting from interspecies ZAR1-ZRK combinations will be critical for expanding known ZRK recognition pathways outside of model systems. Here, we generate a ZED1 mutant library and present an in-depth molecular analysis of ZED1-mediated recognition of HopZ1a through AtZAR1 and NbZAR1. We identify key regions of ZED1 required for HopZ1a recognition, AtZAR1 interaction, and NbZAR1 autoactivity. Finally, we use our ZED1 mutant library to expand HopZ1a recognition to *Solanum lycopersicum* ZAR1 (hereafter SlZAR1).

## Results

### *Error-prone screen reveals new determinants in ZED1 for the recognition of HopZ1a*

To further characterize the molecular functions of ZED1 in HopZ1a recognition, we created a mutant ZED1 library using error-prone PCR amplification [25]. We cloned ZED1 mutants into a dexamethasone-inducible binary vector in frame with a 3' YFP tag (Figure 1A). We removed clones containing only synonymous mutations and/or stop codons. The final library contained 164 clones with an average of 2.8 mutations per clone, covering 72% of ZED1 protein (Figure 1A, Figure S1). We previously demonstrated that the transient co-expression of ZED1 and dexamethasone-inducible HopZ1a in *N. benthamiana* induces a rapid and strong HR that is dependent on *NbZAR1* [7]. We therefore screened each ZED1 mutant for its ability to mediate the recognition of HopZ1a in *N. benthamiana*. We co-expressed each ZED1 clone with HopZ1a and observed HR phenotypes 24 hours after dexamethasone induction (Figure 1B). For each clone, we averaged the score from 6 plants. We identified 85 clones that still retained HopZ1a recognition (referred to as ETI-sufficient), 2 clones with intermediate or variable phenotypes (referred to as ETI-variable), and 77 clones lacking HopZ1a recognition (referred to as ETI-

deficient) (Figure 1C). To ensure that the lack of HR was not due to poor ZED1 expression, we checked for protein expression by western blot analysis (Figure S2, Supplemental Table S1). From 77 ETI-deficient clones, 43 showed protein accumulation similar to ZED1$^{WT}$, 7 showed weak protein accumulation, and 27 had no or low protein accumulation. Therefore, 50 ETI-deficient clones were included for further analysis (Figure S2, Supplemental Table S1).

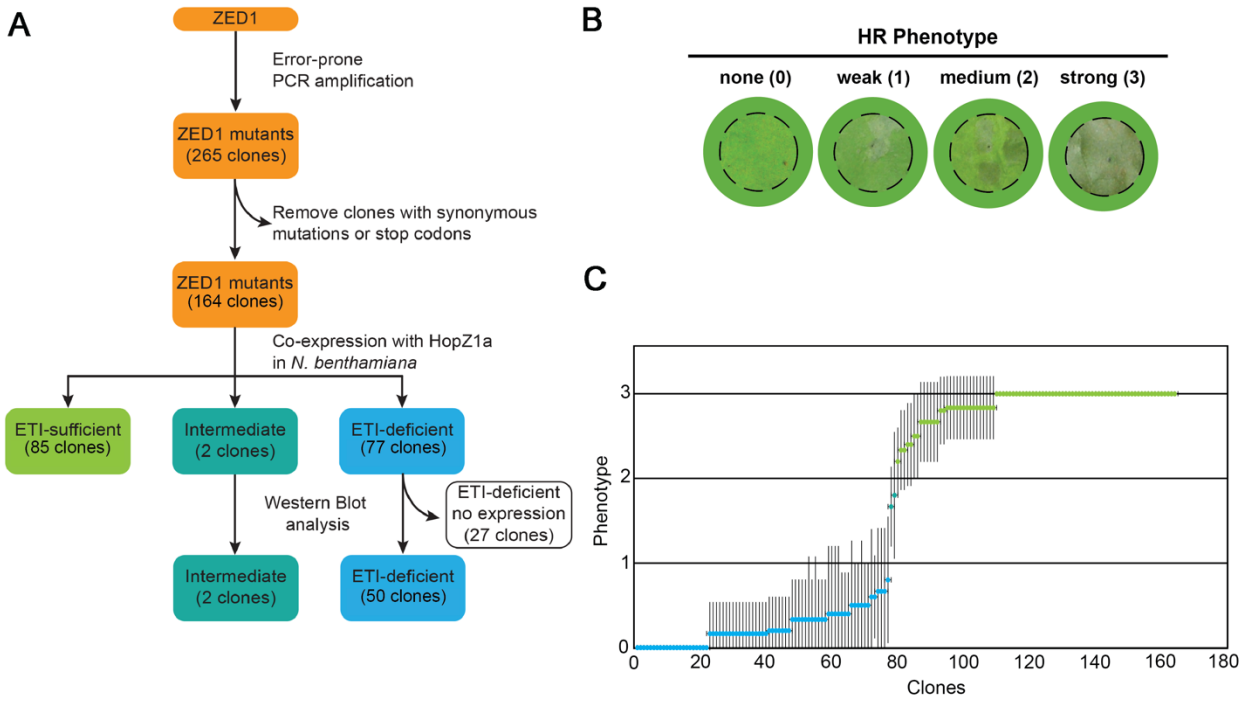



**Figure 1**. A ZED1 error-prone PCR library screen identifies 50 mutants that result in ETI-deficiency. (A) Flow chart depicting the generation of a ZED1 error-prone PCR library, resulting in 164 full length ZED1 mutants. All mutants were screened for the ability to initiate ETI when transiently expressed with HopZ1a in *N. benthamiana*. For all ETI-deficient mutants, ZED1 expression was tested by western blot, and mutants that produced protein (Figure S2) were used in further analysis. (B) The HR was used to determine ETI-deficiency and ETI-sufficiency based on a 4-point scale as previously described [14]. (C) The phenotype of each ZED1 mutant is depicted with standard deviation (n=5-6 plants). Each ZED1 mutant was co-expressed with HopZ1a in *N. benthamiana*, and the HR was scored on a 4-point scale 24 hours after dexamethasone induction.

As most of the clones carried multiple mutations (Figure S1), we sought to identify which mutations were likely responsible for the ETI-deficient phenotype. We separately mapped the mutations from ETI-sufficient clones and ETI-deficient clones, and calculated the number of times each residue was mutated (Figure S3). In addition, we generated a homology based structural model of ZED1 in complex with AtZAR1 based on the AtZAR1-RKS1 inactive conformation structure (Figure S4; PDB:6j5w; Wang, Hu, *et al.*, 2019). Since the activation loop of RKS1 is not resolved in the AtZAR1 inactive structure, the activation loop of ZED1 was modeled with low confidence and was removed from the homology model and future analysis. Using our ZED1 homology model, we identified residues in ZED1 predicted to be in contact with AtZAR1. We prioritized a) residues that were mutated in ETI-deficient clones which carried only a single mutation, b) residues that were mutated in at least two independent ETI-deficient clones but not mutated in ETI-sufficient clones, or c) residues that were predicted to be in the

AtZAR1-ZED1 interface. With these parameters, we selected 28 residues across 24 clones that were likely important for ZED1 function (Figure S3).

*Interaction with ZAR1$^{LRR}$ is perturbed in ZED1 mutants*

HopZ1a recognition through ZED1 requires ZED1-AtZAR1$^{LRR}$ interaction. We have shown that mutations in AtZAR1 or ZED1 that abolished this interaction resulted in an ETI-deficient phenotype, and all mutations that are ETI-sufficient have retained ZED1-AtZAR1$^{LRR}$ interaction [7,12,18]. We thus tested whether clones carrying the candidate mutations were affected in their ability to interact with AtZAR1$^{LRR}$. We co-expressed each ZED1 clone carrying an in-frame C-terminal YFP tag with AtZAR1$^{LRR}$ (aa 505-852) carrying an in-frame C-terminal 5x-Myc tag. We conducted co-immunoprecipitation assays to determine which clones retained their ability to interact with AtZAR1$^{LRR}$. We excluded clones that had weak protein expression (Supplemental Table S1, Figure S2). From 22 tested clones, we were not consistently able to detect protein levels similar to ZED1$^{WT}$ during co-IP experiments for 5 clones (Figure S5), which were therefore excluded from further analysis. Incredibly, from the remaining 17 clones, only 2 retained the ability to interact with AtZAR1$^{LRR}$ (Table 1; Figure 2). The 2 ETI-deficient clones that retained their ability to interact with AtZAR1$^{LRR}$ were mutated in a known HopZ1a acetylation site (ZED1$^{T177M}$) or the ZED1$^{DFG}$ kinase motif (ZED1$^{W193R}$). We also identified AtZAR1$^{LRR}$ interaction-deficient clones with mutations in other ZED1 kinase motifs; W63 is in the glycine-rich loop, V170 is adjacent to the HRD motif, and T177 and F180 are in the catalytic loop (Figure S3).

**Table 1.** Prioritized ETI-deficient ZED1 clones and candidate mutations based on selection criteria

| | Selection criteria | | | | |
|---|---|---|---|---|---|
| **Clone** | **Single mutation** | **Predicted ZAR1 contact*** | **In two or more clones** | **Other mutations** | **AtZAR1$^{LRR}$ interaction** |
| 170 | G110D | | | none | **No** |
| 2 | T177M | | | none | **Strong** |
| 7 | W193R | | | none | **Strong** |
| 129 | **V170G** | **V170G**† | | none | **No** |
| 190 | | E21D / T168I† | | D207N / T223S / I232M / G236D | **No** |
| 240 | | T31A | | T125I / N283D | **No** |
| 137 | | I34N | | S74L / C112Y / S137T | **No** |
| 122 | | H102Q | | D182E / F194S / F215L / E266A / V317G / E331G | **No** |
| 285 | | K103Q | | K76T / N283I | **No** |
| 260 | | K328E | | S157T / D262G / E265G / R310S | **No** |
| 195 | | D12E / E70V | E266D / R310I | M99K / P143R / S234R | **No** |
| 322 | | E22V† / **N32K** | **N32K** | F27L / S57C / S84T / R310K | **No** |
| 42 | | **N32D** | **N32D** / R147G | S85T / L203M / V226F / D288G | **No** |
| 277 | | **K325E** | W144C / **K325E** | I198L / R246G | **No** |
| 309 | | | R10G / R147G | none | **No** |
| 158 | | | R10K / W144R | M268R | **No** |
| 43 | | | W63R / V249E | I154V / G287R / K313M | **No** |
| 224 | | | F180S | E153K / N184Y / D217E | **No** |

†indicates ZAR1$^{NBD}$ contact, bold text indicates mutation fit 2 out of 3 selection criteria

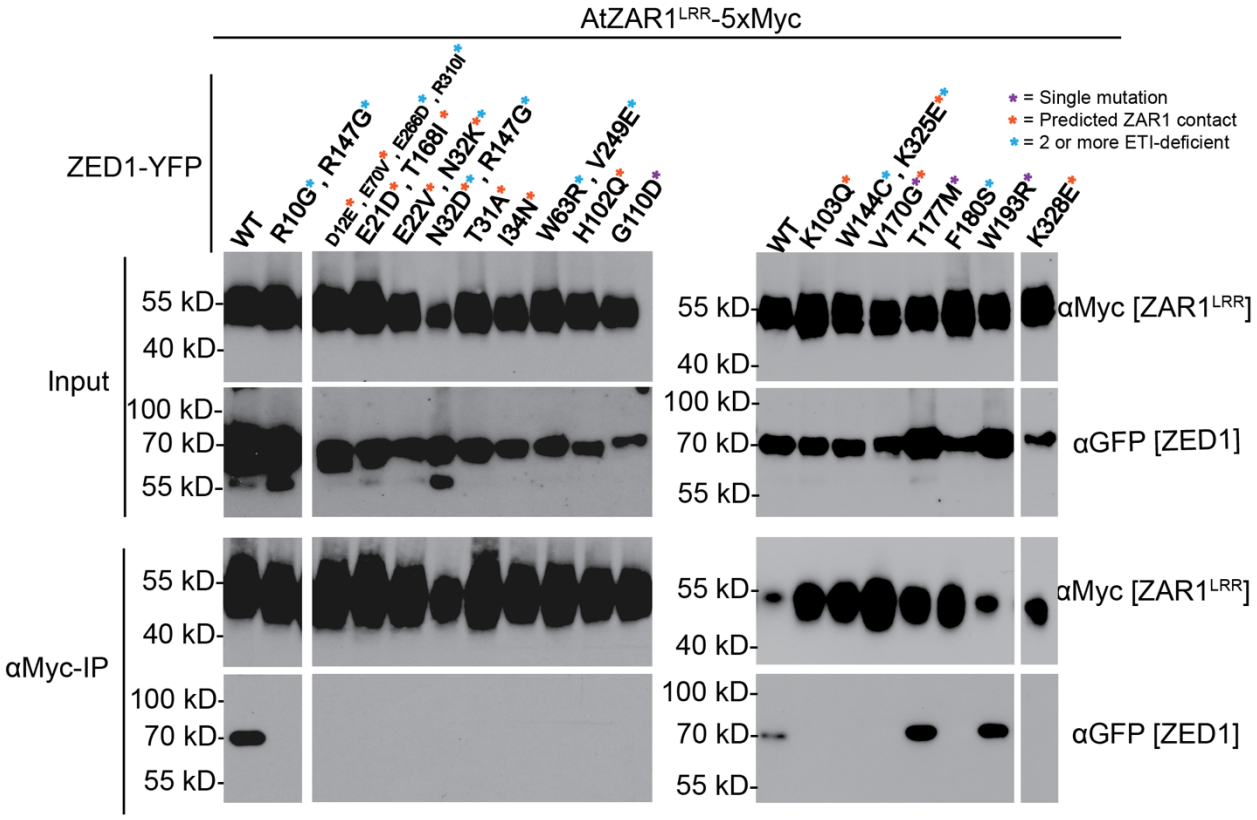
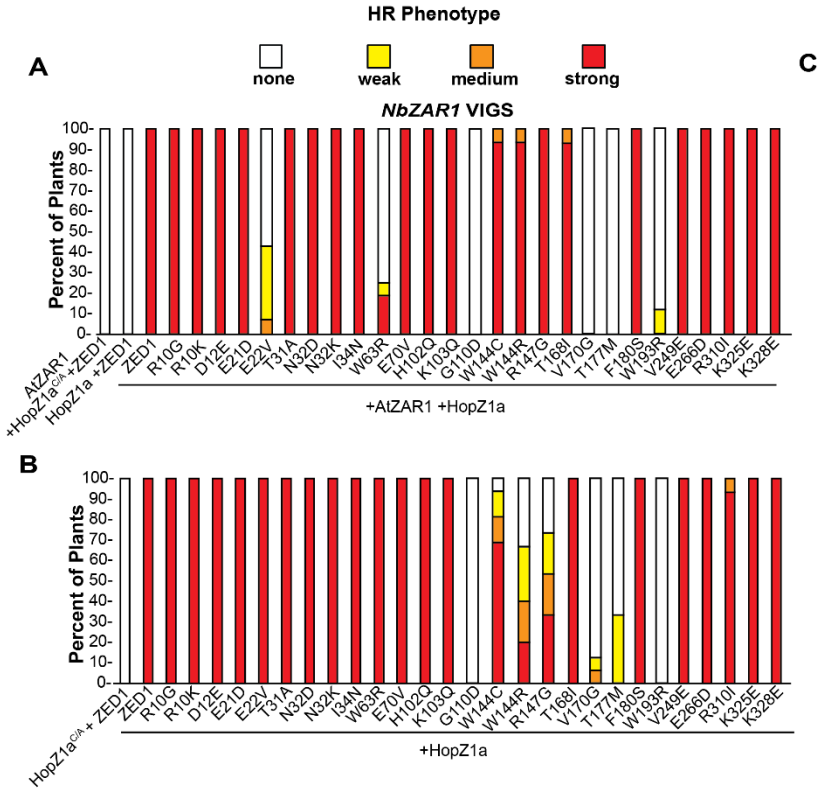



**Figure 2**. ETI-deficient ZED1 mutants show a loss of interaction with AtZAR1$^{LRR}$. Representative western blots of AtZAR1$^{LRR}$ coimmunopurification with ZED1 mutants in *N. benthamiana*. Each ZED1 mutant was cloned as an in-frame fusion with a C-terminal YFP tag in a dexamethasone-inducible expression vector and co-expressed with AtZAR1$^{LRR}$ cloned as an in-frame fusion with a C-terminal 5x-Myc tag. AtZAR1$^{LRR}$ was immunoprecipitated using anti-Myc beads. Colored asterisks indicate the category that prioritized a specific mutation, with purple indicating point mutants, orange indicating predicted ZAR1 contacts, and blue indicating that the mutant was present in 2 or more ETI-deficient clones. Clones may have additional non-candidate mutations not indicated in the labels (see Table 1). Experiments were repeated three times with similar results. Predicted molecular masses of the proteins are AtZAR1$^{LRR}$-5xMyc 50kDa; ZED1-YFP 67 kDa.

*Several ZED1 variants with single mutations are ETI-deficient*

To generate a list of candidate single mutations, we prioritized mutations as described above (Figure S3, Table 1). To investigate the effects of our 27 candidate mutations across 24 ZED1 residues (Table 1), we generated point mutants carrying identical substitutions to those identified in our error-prone screen. We co-expressed each ZED1 mutant, as an in-frame fusion with a 3x Flag tag, with AtZAR1 and HopZ1a in a *NbZAR1*-VIGS background and screened for the HR. Surprisingly, out of 27 candidate point mutants, only ZED1$^{E22V}$, ZED1$^{W63R}$, ZED1$^{G110D}$, ZED1$^{V170G}$, ZED1$^{T177M}$ and ZED1$^{W193R}$ were ETI-deficient (Figure 3A). To determine whether ETI-deficiency was specific to AtZAR1, we co-expressed each ZED1 mutant with HopZ1a in *N. benthamiana*, without silencing the NbZAR1 ortholog [7]. We identified 4 mutants, ZED1$^{G110D}$, ZED1$^{V170G}$, ZED1$^{T177M}$ and ZED1$^{W193R}$, that were ETI-deficient with NbZAR1 (Figure 3B, C).

These four mutations were also ETI-deficient with AtZAR1 (Figure 3A, C). Interestingly, ZED1$^{E22V}$ and ZED1$^{W63R}$ were ETI-deficient with AtZAR1 but ETI-sufficient with NbZAR1. In addition, ZED1$^{W144C}$, ZED1$^{W144R}$ and ZED1$^{R147G}$ were ETI-variable with NbZAR1 but ETI-sufficient with AtZAR1 (Figure 3). These differences indicate that AtZAR1 and NbZAR1 have some functional divergence.




| Mutation | AtZAR1 +HopZ1a HR | Native NbZAR1 +HopZ1a HR |
|---|---|---|
| WT ZED1 | Strong | Strong |
| R10K | Strong | Strong |
| R10G | Strong | Strong |
| D12E | Strong | Strong |
| E21D | Strong | Strong |
| E22V | None | Strong |
| T31A | Strong | Strong |
| N32D | Strong | Strong |
| N32K | Strong | Strong |
| I34N | Strong | Strong |
| W63R | None | Strong |
| E70V | Strong | Strong |
| H102Q | Strong | Strong |
| K103Q | Strong | Strong |
| G110D | None | None |
| W144C | Strong | Variable |
| W144R | Strong | Variable |
| R147G | Strong | Variable |
| T168I | Strong | Strong |
| V170G | None | None |
| T177M | None | None |
| F180S | Strong | Strong |
| W193R | None | None |
| V249E | Strong | Strong |
| E266D | Strong | Strong |
| R310I | Strong | Strong |
| K325E | Strong | Strong |
| K328E | Strong | Strong |

**Figure 3**. ZED1 point mutants show differential HR phenotypes dependent on the ZAR1 ortholog. ZED1 point mutants were cloned into a dexamethasone-inducible expression vector and transiently co-expressed with (A) AtZAR1 and HopZ1a in *N. benthamiana* silenced for *NbZAR1* via virus-induced gene silencing (VIGS) or (B) HopZ1a in *N. benthamiana* [7]. The HR was scored on a 4-point scale (Figure 1B) 24 hours after dexamethasone induction (n=14-17 plants). HopZ1a$^{C216A}$ was co-expressed with ZED1 (A) or AtZAR1 and ZED1 (B) as a negative control. To ensure VIGS of *NbZAR1* was successful, HopZ1a and ZED1 were co-expressed in each *NbZAR1*-VIGS plant [7]. (C) Summary table showing the phenotype of each ZED1 point mutant with either AtZAR1 (A) or NbZAR1 (B). Data are compiled from at least three separate experiments.

Single mutations in ZED1 have previously been shown to cause an autoactive HR phenotype in *A. thaliana* and *N. benthamiana* [16,19]. Therefore, it is possible that our ETI-sufficient mutants caused an HR due to autoactivity instead of HopZ1a recognition. To test whether the ETI-sufficient point mutants resulted in an autoactive HR with NbZAR1, we transiently expressed each ZED1 mutant in *N. benthamiana*. However, we did not observe any autoactive HR phenotypes (Figure S6). Next, to determine if the ZED1 point mutants resulted in an autoactive HR phenotype with AtZAR1, we co-expressed each ETI-sufficient point mutant with AtZAR1 in a *NbZAR1*-VIGS background and screened for an HR. We did not see any autoactive HR phenotypes for any of the point mutants when co-expressed with AtZAR1 (Figure S6).

*ZED1 point mutants have perturbed interaction with AtZAR1$^{LRR}$*

Based on our current understanding of the ZAR1 system, it is most likely that ZED1 mutations leading to ETI-deficiency are caused by either loss of interaction with AtZAR1$^{LRR}$ or HopZ1a [7,12,18]. Of the 6 ETI-deficient point mutants, 4 were tested for their ability to interact with AtZAR1$^{LRR}$ as a part of our initial screen (Figure 2). ZED1$^{G110D}$ and ZED1$^{V170G}$ could not interact with AtZAR1$^{LRR}$, while ZED1$^{T177M}$ and ZED1$^{W193R}$ could interact with AtZAR1$^{LRR}$ (Figure 2). To determine if the two remaining point mutants, ZED1$^{E22V}$ and ZED1$^{W63R}$, were able to interact with AtZAR1$^{LRR}$, we transiently co-expressed each with AtZAR1$^{LRR}$ in *N. benthamiana* and conducted co-IP experiments. We found that ZED1$^{E22V}$ interacted with AtZAR1$^{LRR}$ with similar strength to ZED1$^{WT}$, while ZED1$^{W63R}$ showed a weaker AtZAR1$^{LRR}$ interaction (Figure 4A). Thus, out of the 6 ETI-deficient ZED1 point mutants, 4 mutants (ZED1$^{E22V}$, ZED1$^{W63R}$, ZED1$^{T177M}$, ZED1$^{W193R}$) retain their ability to either fully or partially interact with AtZAR1$^{LRR}$, suggesting that their loss-of-function might instead arise from a compromised interaction with HopZ1a.

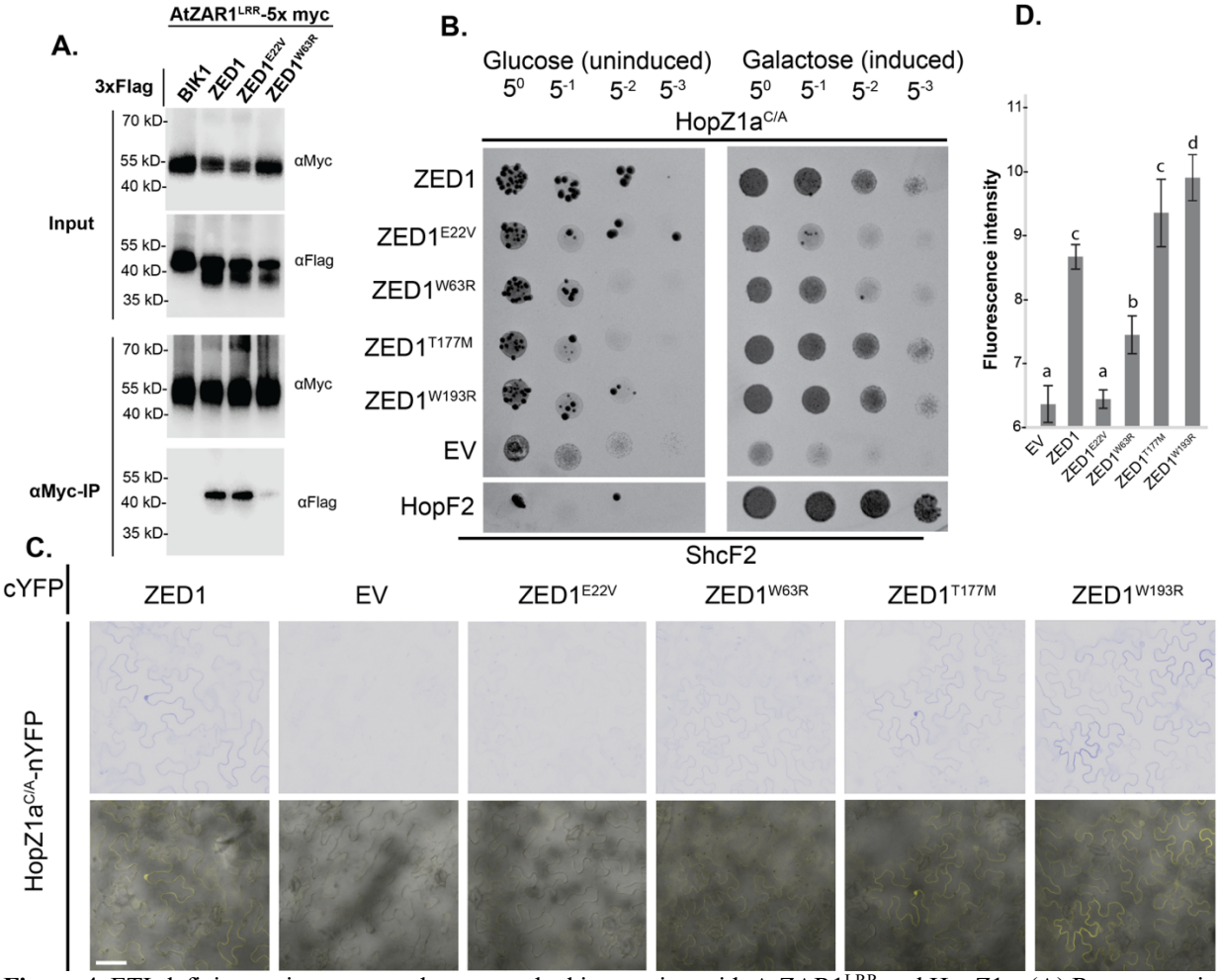



**Figure 4**. ETI-deficient point mutants show perturbed interaction with AtZAR1$^{LRR}$ and HopZ1a. (A) Representative western blots of AtZAR1$^{LRR}$ coimmunopurification with ZED1 mutants in *N. benthamiana.* All ETI-deficient ZED1 point mutants that were not previously tested for ZAR1$^{LRR}$ interactions (see Figure 2) were cloned as in-frame C-terminal fusions with a 3xFlag tag co-expressed with AtZAR1$^{LRR}$ with an in-frame C-terminal 5xMyc tag in *N. benthamiana*. AtZAR1$^{LRR}$ was immunoprecipitated using Anti-Myc beads. (B) ZED1 point mutants were fused in-frame to the DNA-binding domain and tested for interaction with HopZ1a$^{C216A}$ fused in-frame to the activation domain in the LexA yeast two-hybrid assay. As a positive control for interaction, we used HopF2$_{Pto}$DC3000 and its

chaperone ShcF2$_{PtoDC3000}$ that are known to strongly interact [26]. (C) Representative images showing the inverted YFP channel (top) and YFP and brightfield channel merged (bottom) from bimolecular fluorescence complementation (BiFC) assays show perturbed interaction between ZED1 point mutants and HopZ1a$^{C216A}$. All ZED1 point mutants were cloned as in-frame fusions with cYFP, and were co-expressed with HopZ1a$^{C216A}$ fused in-frame with nYFP in *N. benthamiana*. Images were captured 24 hours after dexamethasone induction using a Zeiss LSM710 confocal scanning microscope. White bar indicates 50μm. (D) Mean fluorescence intensity from all fields imaged in BiFC experiments was calculated using the Leica Application Suite X software (n=12-13 images across 6 plants and 3 experiments). Error bars represent standard error of means. T-tests were conducted to determine statistical significance, indicated by letters above bars ($p<0.05$). All experiments were repeated three times with similar results.

To better contextualize our AtZAR1-ZED1 interaction data as well as previous mutations known to perturb AtZAR1-ZED1 interactions, we generated an updated model of ZED1 in complex with AtZAR1 using AlphaFold Multimer [27]. In the cryo-EM structure of RKS1 in complex with inactive AtZAR1, the activation loop of RKS1 is absent. The absence of this key region confounds homology modeling efforts of the ZED1 activation loop. In our AtZAR1-ZED1 AlphaFold Multimer structure, the activation loop of ZED1 is resolved (Figure S7). However, in the AlphaFold model, the ZED1 activation loop is modeled with both contacts and clashes with the AtZAR1$^{NBD}$, resulting in a slight displacement of AtZAR1$^{NBD}$ relative to the cryo-EM structure and our original AtZAR1-ZED1 homology model (Figure S7). This result was surprising, since displacement of the AtZAR1$^{NBD}$ by the RKS1 activation loop triggers AtZAR1 activation, and AtZAR1 and ZED1 autoactivity has not been observed. Next, we mapped all single residues affecting ZED1-AtZAR1 interaction onto our AtZAR1-ZED1 AlphaFold model. This included ZED1$^{G110}$, ZED1$^{V170}$ and ZED1$^{W63}$ identified here, ZED1$^{I24}$ and ZED1$^{G29}$, which are required for ZED1-AtZAR1 interaction [28], and ZED1$^{G128}$ and ZED1$^{D231}$, which are required for ZED1-AtZAR1$^{LRR}$ interaction (Figure S8) [7]. Mutations in I24, G29 or D231 of ZED1 have been shown to weaken or abolish ZED1 mediated HR [7,28], while mutations in G128 of ZED1 weaken AtZAR1$^{LRR}$ interaction but do not result in loss of ZED1-mediated HR [7]. To our surprise, only one residue, ZED1$^{I24}$ is in direct contact with AtZAR1 in our AlphaFold model (Figure S8). The remaining 6 residues are in 3 clusters, in regions potentially involved in maintaining the integrity of the kinase scaffold (Figure S8). G110 and W63 of ZED1 are each in the middle of a central β-strand in a β-sheet while G29 is in contact with I24 in a loop following the N-terminal α-helix (Figure S8). D231 is in an α-helix bundle at the end of the ZED1 activation loop and is adjacent to V170 which is in contact with the activation loop (Figure S8). Finally, G128 is in contact with a region immediately preceding the ZED1 activation loop. This suggests that point mutants need to destabilize protein regions, especially the ZED1 activation loop, to be ETI-deficient.

*ZED1 point mutants show loss of interaction with HopZ1a in yeast and in planta*

We hypothesized that ETI-deficient ZED1 mutants that were able to interact with AtZAR1$^{LRR}$ may have perturbed interaction with HopZ1a. To determine if the 4 AtZAR1$^{LRR}$ interaction-sufficient, ETI-deficient mutants, ZED1$^{E22V}$, ZED1$^{W63R}$, ZED1$^{T177M}$ and ZED1$^{W193R}$, can interact with HopZ1a, we used the LexA yeast-two hybrid system. We cloned each ZED1 point mutant as an in-frame fusion with the activation domain and tested for interaction with the catalytically inactive mutant HopZ1a$^{C216A}$ cloned as an in-frame fusion to the DNA-binding domain. The HopZ1a$^{C216A}$ catalytic mutant has previously been shown to interact strongly with ZED1 [7,12]. We spotted the same optical density of yeast onto plates to better quantify the interaction strength. We noted that ZED1$^{E22V}$ and ZED1$^{W63R}$ showed a weaker interaction with HopZ1a$^{C216A}$ compared to ZED1, ZED1$^{T177M}$ and ZED1$^{W193R}$ (Figure 4B). This weaker

interaction was not due to lack of expression, as the protein expression level of all ZED1 mutants were comparatively more abundant than that of ZED1$^{WT}$ in yeast (Figure S9).

To confirm the observed ZED1-HopZ1a$^{C216A}$ interactions in planta, we used bimolecular fluorescence complementation (BiFC). We cloned ZED1$^{E22V}$, ZED1$^{W63R}$, ZED1$^{T177M}$ and ZED1$^{W193R}$ as a fusion with the C-terminal fragment of YFP and co-expressed each mutant with HopZ1a$^{C216A}$ fused to the N-terminal fragment of YFP. We examined leaf tissue under a confocal microscope and found that ZED1$^{E22V}$ and ZED1$^{W63R}$ interacted less strongly with HopZ1a when compared to ZED1$^{WT}$, as indicated by lower fluorescence intensity (Figure 4C). Interestingly, ZED1$^{T177M}$ and ZED1$^{W193R}$ had a higher mean fluorescence relative to ZED1$^{WT}$, indicating a more robust interaction with AtZAR1, although this difference was only significant for ZED1$^{W193R}$ (Figure 4C, D). We conclude that ZED1$^{E22V}$ is likely ETI-deficient due to its perturbed interaction with HopZ1a, and that ZED1$^{W63R}$ is ETI-deficient due to a weaker interaction with both AtZAR1$^{LRR}$ and HopZ1a, potentially due to broader structural destabilization of ZED1 (Figure S8).

*ZED1 is a pseudokinase, and is unable to bind nucleotides or cations*

Since ZED1$^{T177M}$ and ZED1$^{W193R}$ are ETI-deficient, but can interact with AtZAR1$^{LRR}$ and HopZ1a, we wanted to further investigate why these mutations result in ETI-deficiency. It is possible that ZED1$^{T177M}$ results in ETI-deficiency because it is in a known HopZ1a acetylation site [12]. However, we have previously shown that the ZED1$^{T177A}$ mutation is ETI-sufficient [7]. The ZED1$^{W193R}$ mutation occurs in the DFG motif of the ZED1 kinase-like domain, which is responsible for $Mg^{2+}$ cofactor binding in conventional kinases, while the ZED1$^{T177M}$ mutation is in the catalytic loop near the HRD motif (Figure S3). We have previously shown that ZED1 lacks kinase activity [12]. However, we hypothesized that ZED1 may retain some nucleotide or cation binding activity important for HopZ1a recognition, and that this binding may be inhibited in ZED1$^{T177M}$ and ZED1$^{W193R}$. To test this hypothesis, we conducted an in-vitro thermal shift assay [29]. We tested $Mg^{2+}$ and nucleotides known to be bound by kinases (ATP, ADP, AMP, AMPPNP), as well as two promiscuous kinase inhibitors (Vi16832, DAP) [30]. ZED1 did not detectably bind nucleotides, cations, or pan-kinase inhibitors in our assay (Figure S10). Since neither ZED1$^{T177M}$ or ZED1$^{W193R}$ restored a kinase motif synonymous with conventional protein kinase catalytic activity, we did not assay these mutants.

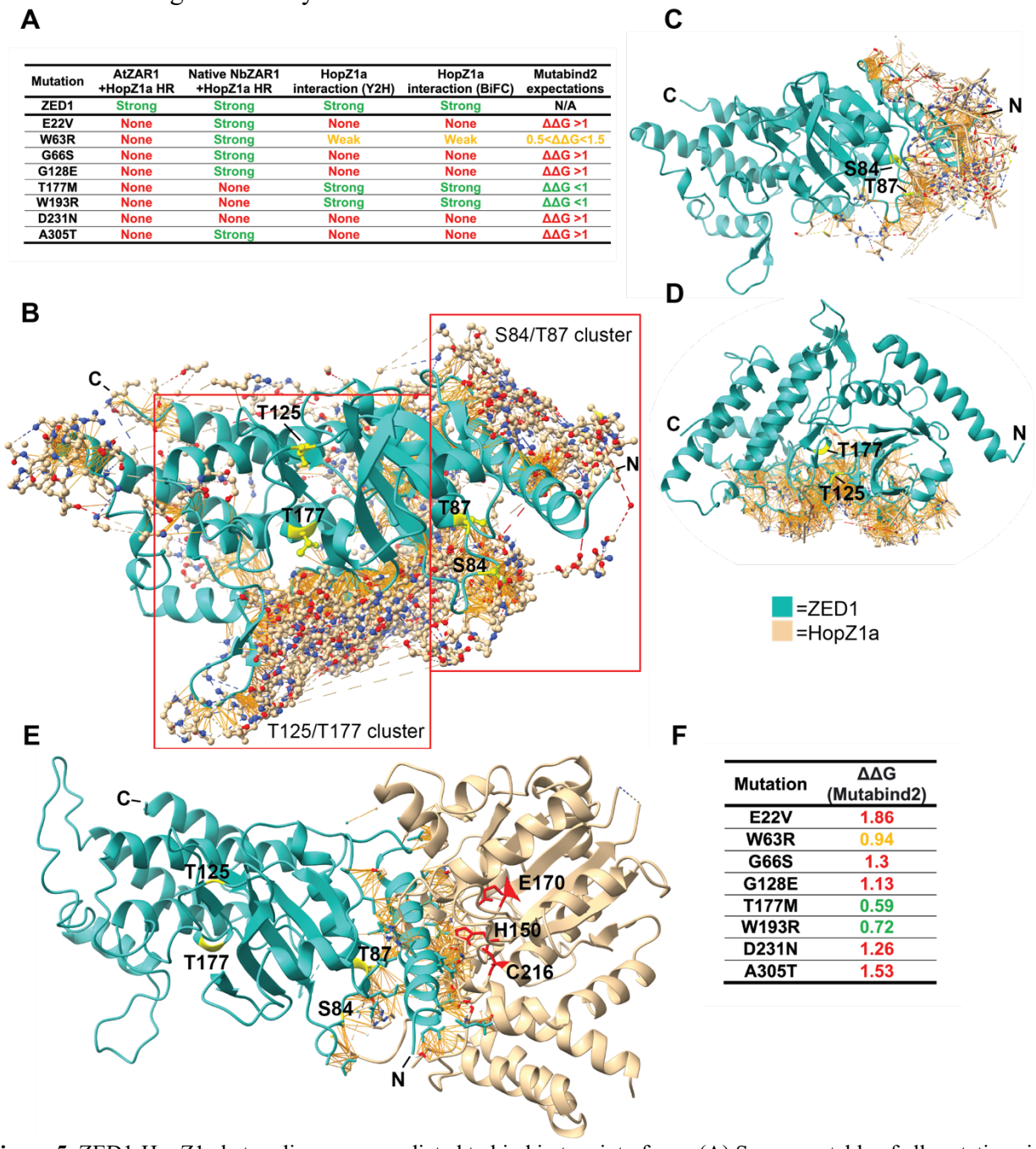
*Refined docking analysis supports the existence of two ZED1-HopZ1a binding sites*

To conduct an informed ZED1-HopZ1a docking analysis, we considered our ZED1-HopZ1a interaction data, along with previously published data on ZED1-HopZ1a interactions [7], the HopZ1a crystal structure [31], our ZED1 homology model (Figure S4), and our ZED1 AlphaFold model (Figure S7). We used Cluspro, Hdock, and LZerD to generate ZED1-HopZ1a heterodimer models without constraints on specific receptor or ligand regions [32–34]. We ran these programs using either the ZED1 homology model, or our AlphaFold model. We selected the top model from each program and used Mutabind2 to predict the change in binding affinity resulting from all known ZED1 mutations at 8 amino acids with experimental ZED1-HopZ1a interaction data [35] (Figure 5A, Supplemental Table S2). Mutabind2 binding affinity ($\Delta\Delta G$) predictions fit experimental data if a) $\Delta\Delta G > 1$ for mutations resulting in loss of ZED1-HopZ1a interactions, b) $\Delta\Delta G < 1$ for mutations with no effect on ZED1-HopZ1a interactions and c) $1.5 > \Delta\Delta G < 0.5$ for mutations resulting in weaker ZED1-HopZ1a interactions (Figure 5A). We prioritized models where Mutabind2 predictions for at least 7 out of 8 amino acids fit experimental ZED1-HopZ1a

interaction data (Figure 5A). From 80 evaluated models, 20 fit Mutabind2 expectations for 7 out of 8 amino acids, with 2 models satisfying all 8 amino acid criteria (Supplemental Table S2). Interestingly, 17 of the 20 models fitting our experimental predictions identify two clusters. One cluster surrounds the N-terminus of ZED1 near acetylation sites S84 and T87 while the second cluster surrounds the activation loop near T125 and T177 (Figure 5B). This suggests that ZED1 may have two HopZ1a binding interfaces, one resulting in the acetylation of S84 and T87, and the other resulting in the acetylation of T125 and T177.

A

| Mutation | AtZAR1 +HopZ1a HR | Native NbZAR1 +HopZ1a HR | HopZ1a interaction (Y2H) | HopZ1a interaction (BiFC) | Mutabind2 expectations |
|---|---|---|---|---|---|
| ZED1 | Strong | Strong | Strong | Strong | N/A |
| E22V | None | Strong | None | None | ΔΔG >1 |
| W63R | None | Strong | Weak | Weak | 0.5<ΔΔG<1.5 |
| G66S | None | Strong | None | None | ΔΔG >1 |
| G128E | None | Strong | None | None | ΔΔG >1 |
| T177M | None | None | Strong | Strong | ΔΔG <1 |
| W193R | None | None | Strong | Strong | ΔΔG <1 |
| D231N | None | None | None | None | ΔΔG >1 |
| A305T | None | Strong | None | None | ΔΔG >1 |




F

| Mutation | ΔΔG (Mutabind2) |
|---|---|
| E22V | 1.86 |
| W63R | 0.94 |
| G66S | 1.3 |
| G128E | 1.13 |
| T177M | 0.59 |
| W193R | 0.72 |
| D231N | 1.26 |
| A305T | 1.53 |

**Figure 5**. ZED1-HopZ1a heterodimers are predicted to bind in two interfaces. (A) Summary table of all mutations in ZED1 that have been shown to affect HopZ1a interaction, the phenotype of these mutations, and expectations for ZED1-HopZ1a heterodimer models assessed in Mutabind2 based on experimental data [7,35]. Mutabind2 expectations are shown as the change in binding affinity (ΔΔG), with positive values indicated a decrease in binding affinity, and

values closer to zero indicating no change on binding affinity. (B) ZED1 AlphaFold model (blue) and atoms from HopZ1a (beige) that are in contact (orange lines) with ZED1 from top unconstrained docked models produced using Cluspro, Hdock, and LZerD [32–34] that meet Mutabind2 expectations for 7 out of 8 amino acids. Models that meet experimental expectations form two clusters: S84/T87 cluster around the N-terminus of ZED1, and T125/T177 cluster surrounding the activation loop (both clusters are boxed in red). Docking analysis was constrained to the two putative HopZ1a binding interfaces, where HopZ1a's catalytic triad (H150, E170, and C216) was in contact with either (C) S84/T87 acetylation sites, or (D) T125/T177 acetylation sites. Depictions of constrained models show top models (Mutabind2 expectations for 7 out of 8 amino acids) with contacts (orange lines) between HopZ1a interface atoms (beige) and the ZED1 AlphaFold model (blue). (E) The most likely ZED1-HopZ1a heterodimer model based on experimental data. The ZED1 homology model (blue) and acetylated sites (yellow), as well as HopZ1a (beige) and the catalytic site (red) are in contact (orange lines) primarily with the ZED1 around the N-terminal α-helix of ZED1 near the S84/T87 acetylation site. (F) The Mutabind2 predictions for our proposed heterodimeric ZED1-HopZ1a model fit expectations based on experimental data.

To further explore our models for the two putative HopZ1a binding interfaces, we employed LzerD and Hdock to generate models where the catalytic triad of HopZ1a (H150, E170, and C216) was in contact with either S84/T87 acetylation sites, or T125/T177 acetylation sites. We conducted this analysis with our homology model as well as our AlphaFold model and selected the top 10 models from each approach. We evaluated models using Mutabind2 with the same criteria as for unconstrained models. From 80 evaluated models, 22 fit Mutabind2 predictions for 7 out of 8 amino acids (Figure 5C, D, Supplemental Table S2).

Across all our approaches, only one model fit Mutabind2 predictions for 8 out of 8 amino acids based on our experimental data and positioned the HopZ1a catalytic site near a ZED1 acetylation site cluster (Figure 5E, F). The model was created based on the ZED1 homology model and positions the catalytic site of HopZ1a around the N-terminal α-helix of ZED1 near the S84/T87 acetylation site (Figure 5F). Based on our evaluation strategy, we feel that this is the most probable model of ZED1 in conformation with HopZ1a. Several models fit Mutabind2 predictions for 7 out of 8 amino acids based on our experimental data and placed the HopZ1a active site in proximity with the T125/T177 acetylation sites (Figure 5B, D). Interestingly, over 85% of models constrained to T125/T177 did not fit our experimental expectations regarding the ZED1$^{E22V}$ mutation, suggesting that ZED1$^{E22}$ may not be important for ZED1-HopZ1a interaction at this interface (Supplemental Table S2). ZED1$^{E22V}$ resulted in a loss of HopZ1a interaction but was still ETI-sufficient with NbZAR1 (Figure 4A, Figure 3). It is possible that ZED1 has multiple HopZ1a binding interfaces, with some occurring more transiently than others. Therefore, our experimental ZED1-HopZ1a interaction data may identify the effects of one of these binding interfaces preferentially over another. Furthermore, our data suggests that not all binding conformations are required for ZED1-mediated activation of NbZAR1. We cannot rule out the possibility that ZED1 has intermediate conformations during HopZ1a acetylation and ETI activation which are not captured by our ZED1-HopZ1a docking analysis.

Since ZED1$^{E22V}$ is predicted to primarily impact HopZ1a interaction with the S84/T87 acetylation site, and ZED1$^{E22V}$ is ETI-deficient with AtZAR1 but ETI-sufficient with NbZAR1, we tested whether the T125/T177 and S84/T87 sites are required for NbZAR1 or AtZAR1 mediated recognition of HopZ1a. We have previously shown that ZED1$^{T125A/T177A}$ is ETI-sufficient with NbZAR1 [7], and that this mutant has reduced acetylation by HopZ1a [12]. We generated ZED1$^{S84A/T87A}$ and transiently expressed either this mutant or ZED1$^{T125A/T177A}$ with HopZ1a, or HopZ1a$^{C216A}$ in *N. benthamiana*. We noted a strong HR when ZED1$^{S84A/T87A}$ and ZED1$^{T125A/T177A}$ were co-expressed with HopZ1a but not HopZ1a$^{C216A}$ (Figure S11). To test if mutation of ZED1 acetylation sites impacted AtZAR1-mediated HR, we transiently coexpressed

either ZED1$^{S84A/T87A}$ or ZED1$^{T125A/T177A}$ with HopZ1a, or HopZ1a$^{C216A}$ and AtZAR1 in *NbZAR1*-VIGS plants. While ZED1 $^{S84A/T87A}$ was ETI-sufficient with AtZAR1 and HopZ1a, ZED1$^{T125A/T177A}$ was ETI-deficient (Figure S11). These results further support that HopZ1a acetylates ZED1 at T125 and T177 in planta to trigger AtZAR1-mediated ETI, while ETI in *N. benthamiana* has relaxed requirements for acetylation.

*ZED1 point mutants cause loss of NbZAR1 autoactivity*

We and others have previously shown that ZED1 causes an autoactive HR when transiently co-expressed with NbZAR1 [9,14]. We sought to use our mutant library to gain insight into the mechanisms of autoactive HR resulting from interspecies ZRK-ZAR1 pairings. We transiently co-expressed NbZAR1 with each ZED1 point mutant. From our pool of 27 point mutants, 7 mutants were no longer autoactive, while an additional 5 mutants resulted in variable autoactivity (Figure 6A). We hypothesized that the lack of autoactivity may be due to lower protein accumulation of some ZED1 mutants. However, all ZED1 point mutants that were not autoactive or were variably autoactive were expressed comparably to ZED1$^{WT}$ (Figure 2, Figure 4A, Figure 6B).

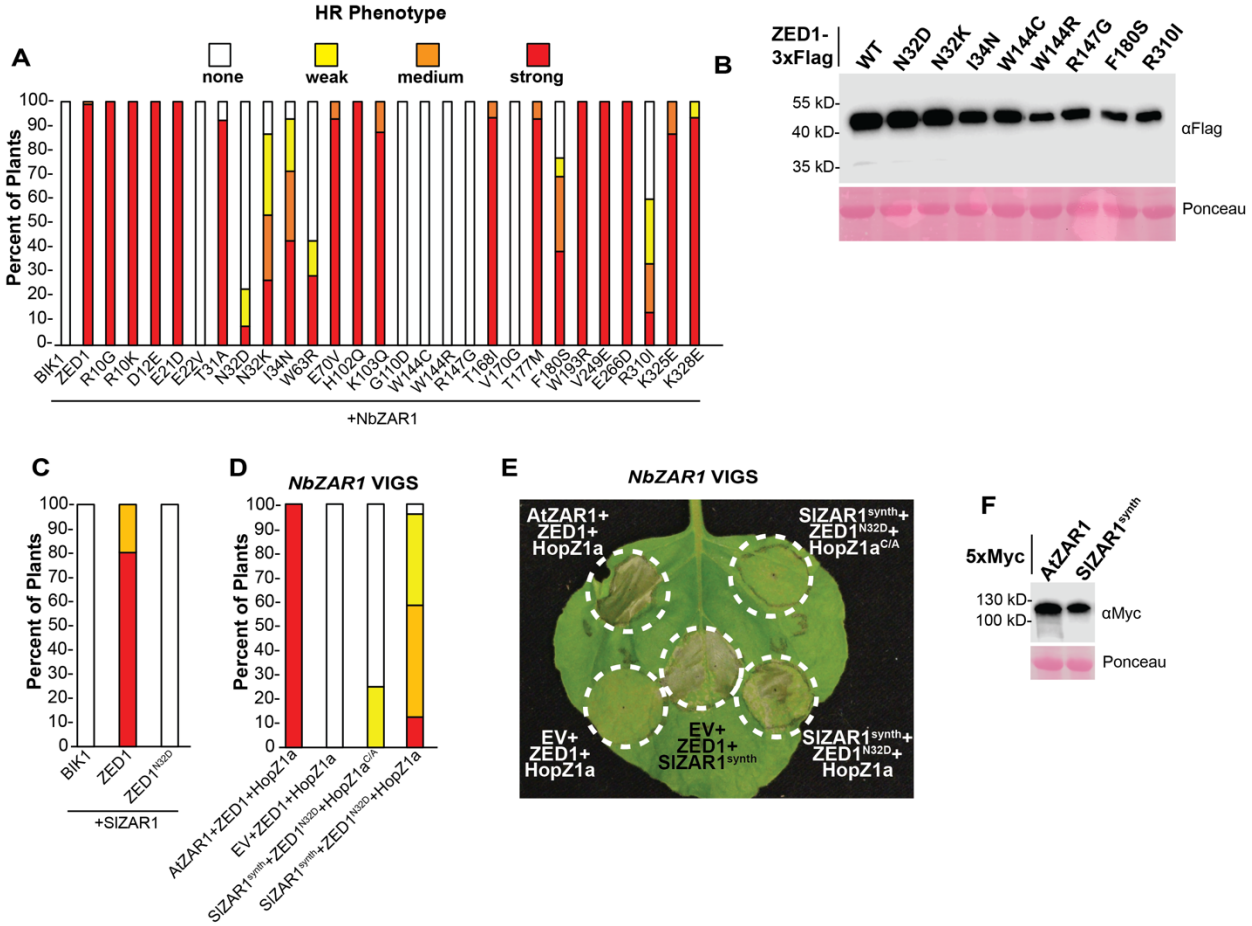



**Figure 6**. ZED1 point mutant library allows for expansion of HopZ1a recognition to SlZAR1. (A) ZED1 point mutants show loss of autoactivity when transiently co-expressed with NbZAR1 in *N. benthamiana*. ZED1 point mutants were cloned with an in-frame C-terminal 3xFlag tag. All constructs were cloned into dexamethasone-inducible vectors and the HR was scored on a 4-point scale (Figure 1B), 24 hours after dexamethasone induction (n=13-17 plants). BIK1 was co-expressed with NbZAR1 as a negative control [14]. (B) Anti-Flag western blot

showing that all ZED1 point mutants that affect NbZAR1 autoactivity are expressed in *N. benthamiana* (see also Figure 2, Figure 4A). (C) Transiently co-expressing SlZAR1 with ZED1, but not BIK1 or ZED1$^{N32D}$ in *N. benthamiana* results in autoactive cell death 48 hours after dexamethasone induction (n=15). (D) ZED1$^{N32D}$ enables recognition of HopZ1a through SlZAR1, and is dependent on HopZ1a catalytic activity, 96 hours post dexamethasone induction in *N. benthamiana NbZAR1* VIGS (n=24). A codon-switched *SlZAR1* construct (*SlZAR1*$^{synth}$) was created to avoid VIGS silencing (see Figure S13). (E) Representative leaf images showing cell death in 4 separate *N. benthamiana* plants 96 hours post-dexamethasone induction. (F) Anti-Myc western blot showing that SlZAR1$^{synth}$ protein from transient expression in *NbZAR1* VIGS plants. Data for all HR assays are compiled from at least three separate experiments.

The autoactive deficient residues concentrated in four distinct clusters in the ZED1 AlphaFold protein structure (Figure S12). At the N-terminus of ZED1, E22, N32, I34 and G110 form a cluster (cluster 1), potentially stabilizing the first α-helix of ZED1, that is near ZAR1$^{NBD}$. In the center of the protein (cluster 2), W63 and F180 contact each other, and a β-strand preceeding the activation loop. (Figure S12). At the C-terminal region of ZED1 (cluster 3), W144 and R147 are in contact with each other and an α-helix at the C-terminal end of the ZED1 activation loop [21,22]. Two additional mutations, ZED1$^{N173S}$ in the HRD motif and ZED1$^{V212T}$ in the activation loop, were previously shown to result in an autoactive phenotype in the absence of HopZ1a [16,19]. Interestingly in our AlphaFold model, N173 is in contact with V212 (Figure S12). ZED1$^{V170G}$ and ZED1$^{R310I}$ are autoactivity deficient and these two residues are in proximity with N173 and V212. Therefore, V170G, N173, V212, and R310 form cluster 4 (Figure S12). In addition, ZED1$^{R310}$ corresponds to RKS1$^{K324}$, which is at the RKS1-PBL2 interface [21,22]. Taken together, residues that are important for the ZED1-NbZAR1 autoactive phenotype cluster around regions of potential importance for stability of the ZED1 activation loop, proximity of ZED1 to AtZAR1$^{NBD}$, and potential ZED1-PBL interfaces.

Two of the autoactive deficient mutants, ZED1$^{G110D}$ and ZED1$^{V170G}$, did not interact with AtZAR1$^{LRR}$ (Figure 2) and were ETI-deficient (Figure 3), potentially because of protein destabilization. Therefore, loss of interaction with NbZAR1 may explain why ZED1$^{G110D}$ and ZED1$^{V170G}$ are autoactive deficient. Two additional autoactive deficient mutants, ZED1$^{E22V}$ and ZED1$^{W63R}$ were ETI-deficient with AtZAR1 but not NbZAR1, suggesting that mechanisms of activation uniquely required for AtZAR1 may be important for NbZAR1 autoactivation. Finally, ZED1$^{N32D}$ was the only mutant that was not autoactive, and was ETI-sufficient with both AtZAR1 and NbZAR1 (Table 2). An ETI-sufficient, autoactive-deficient mutant such as ZED1$^{N32D}$ presents an opportunity for expansion of ZED1-mediated HopZ1a recognition outside of model species.

**Table 2.** Summary of ZED1 point mutant phenotypes and interaction data

| Mutation | AtZAR1 +HopZ1a HR | Native NbZAR1 +HopZ1a HR | NbZAR1 Autoactive HR | Native NbZAR1 Autoactive HR | AtZAR1 Autoactive HR | Original Selection Category* | HopZ1a interaction | AtZAR1$^{LRR}$ interaction |
|---|---|---|---|---|---|---|---|---|
| WT ZED1 | Strong | Strong | Strong | None | None | | Strong | Strong |
| R10K | Strong | Strong | Strong | None | None | 2+ | | |
| R10G | Strong | Strong | Strong | None | None | 2+ | | |
| D12E | Strong | Strong | Strong | None | None | Contact | | |
| E21D | Strong | Strong | Strong | None | None | Contact | | |
| E22V | None | Strong | None | None | None | Contact | None | Strong |
| T31A | Strong | Strong | Strong | None | None | Contact | | |
| N32D | Strong | Strong | None | None | None | Contact, 2+ | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N32K | Strong | Strong | Variable | None | None | Contact, 2+ | | |
| I34N | Strong | Strong | Variable | None | None | Contact | | |
| W63R | None | Strong | None | None | None | 2+ | Weak | Weak |
| E70V | Strong | Strong | Strong | None | None | Contact | | |
| H102Q | Strong | Strong | Strong | None | None | Contact | | |
| K103Q | Strong | Strong | Strong | None | None | Contact | | |
| G110D | None | None | None | None | None | Single | | None |
| W144C | Strong | Variable | None | None | None | 2+ | | |
| W144R | Strong | Variable | None | None | None | 2+ | | |
| R147G | Strong | Variable | None | None | None | 2+ | | |
| T168I | Strong | Strong | Strong | None | None | Contact | | |
| V170G | None | None | None | None | None | Single, Contact | | None |
| T177M | None | None | Strong | None | None | Single | Strong | Strong |
| F180S | Strong | Strong | Variable | None | None | 2+ | | |
| W193R | None | None | Strong | None | None | Single | Strong | Strong |
| V249E | Strong | Strong | Strong | None | None | 2+ | | |
| E266D | Strong | Strong | Strong | None | None | 2+ | | |
| R310I | Strong | Strong | Variable | None | None | 2+ | | |
| K325E | Strong | Strong | Strong | None | None | Contact, 2+ | | |
| K328E | Strong | Strong | Strong | None | None | Contact | | |

*Selection criteria refer to the reason the mutation was selected as a candidate (see Table 1). Single indicates the mutation was a single mutation in the original screen, 2+ indicates the mutation was in two or more ETI-deficient clones, and contact means the residue is predicted to be in the ZAR1-ZED1 interface.

*ETI-sufficient, NbZAR1 autoactive-deficient ZED1$^{N32D}$ mutant allows for SlZAR1 recognition of HopZ1a*

ZAR1-ZRK incompatibility between species needs to be overcome before ZRK effector recognition pathways can be translated into non-model systems. We previously attempted to generate transgenic *Solanum lycopersicum* expressing *A. thaliana* native promoter-driven *ZED1* but were unable to recover viable seeds from transformants expressing the transgene (data not shown). To determine if ZED1 leads to a *Solanum lycopersicum* ZAR1 (hereafter SlZAR1) dependent HR phenotype in the absence of pathogens, we co-expressed ZED1 with SlZAR1 in *N. benthamiana*. We noted a strong HR phenotype when SlZAR1 was co-expressed with ZED1, but not when SlZAR1 was co-expressed with BIK1 (Figure 6C). However, this HR was slightly slower than the NbZAR1 autoactive phenotype, and HRs were scored 48 hours after dexamethasone induction, rather than 24 hours.

Next, we wanted to determine if one of the ZED1 mutants could abolish the autoactive SlZAR1 phenotype, while enabling HopZ1a recognition through SlZAR1. We selected ZED1$^{N32D}$ as it was the only mutation that could mediate recognition of HopZ1a through both NbZAR1 and AtZAR1, and was autoactive-deficient with NbZAR1 (Figure 6A, Table 2). We co-expressed ZED1$^{N32D}$ with SlZAR1 in *N. benthamiana* and noted a loss of the autoactive HR phenotype (Figure 6C). To determine if HopZ1a recognition through SlZAR1 could be achieved with ZED1$^{N32D}$, we generated a *SlZAR1* construct with altered codons (hereafter *SlZAR1$^{synth}$*) to prevent silencing in an *NbZAR1*-VIGS background (Figure S13), and co-expressed SlZAR1$^{synth}$ with ZED1$^{N32D}$ and HopZ1a. We noticed a strong and consistent HR phenotype when SlZAR1$^{synth}$ was co-expressed with HopZ1a, but not HopZ1a$^{C216A}$, indicating that the catalytic

activity of HopZ1a was necessary for HR. However, the observed HR was delayed relative to co-expression of AtZAR1, ZED1 and HopZ1a, occurring 96 hours post-induction (Figure 6D, E). This delayed phenotype was not due to lower SlZAR1$^{synth}$ expression relative to AtZAR1 (Figure 6F). These data show that SlZAR1 can recognize modification of ZED1$^{N32D}$ by HopZ1a, although it is not as efficient as AtZAR1 or NbZAR1 in our transient assay. To confirm that the ZED1$^{N32D}$ mutation was not enabling HopZ1a recognition through a different NLR in *N. benthamiana*, we co-expressed ZED1$^{N32D}$ with HopZ1a in *NbZAR1*-VIGS plants and did not observe an HR after 96 hours (Figure S14).

**Discussion**

Here, we identify key determinants underlying the molecular mechanisms of ZED1-HopZ1a interaction, ZED1-AtZAR1 interaction, and ZED1-ZAR1 interspecies autoactive cell death. We show that ZED1-mediated immunity in response to HopZ1a and AtZAR1-ZED1 interaction are perturbed by higher order mutations (Figure 1, Figure 2) and are highly resilient to single mutations (Figure 2, Figure 3, Figure 4A). We tested four ETI-deficient point mutants that can interact with AtZAR1 for their ability to interact with HopZ1a (Figure 4) and used this and previous data to identify two binding interfaces for HopZ1a that fit our experimental data (Figure 5). Finally, using our library of ZED1 point mutants, we identified four clusters of residues that underpin interspecies ZAR1-ZRK phenotypes (Figure 6, Supplemental Figure S12). We used an ETI-sufficient ZED1 point mutant that does not result in autoactive cell death to achieve SlZAR1-mediated recognition of HopZ1a via transient expression in *N. benthamiana* (Figure 6). This represents the first expansion of a ZRK recognition pathway to a ZAR1 ortholog from a non-model species.

While ZAR1$^{LRR}$ is the key determinant of ZRK interaction specificity, we found that most single mutations in the putative ZED1-AtZAR1$^{LRR}$ interaction interface did not result in ETI-deficiency (Figure 3, Table 2). However, a loss of interaction between AtZAR1$^{LRR}$ and ZED1 was observed for all tested ZED1 mutants carrying multiple mutations (Figure 2). We have previously shown that AtZAR1$^{LRR}$ interactions are sensitive to single mutations in ZED1 and ZRK10 [7,14,18]. We have also shown ZED1 point mutants that are loss-of-function in *A. thaliana* may be ETI-sufficient in *N. benthamiana* [7]. Given that our screen was conducted in *N. benthamiana*, we expected more ZED1 point mutants to be ETI-deficient in *N. benthamiana*. However, we found that ZED1-mediated recognition of HopZ1a is remarkably resilient to single mutations when assayed via transient expression (Table 2). Bacterial effectors and plant ETI components are under strong selective pressure [36–39]. It is possible that the functional resilience of ZED1 to mutagenesis enables evolutionary plasticity, such as that recently described for the primate antiviral protein TRIM5α, which directly binds retroviral capsids [40,41]. When the TRIM5α retroviral capsid binding region is randomly mutated, the majority of mutations do not disrupt viral restriction, and some mutations improve viral restriction [40,41]. This mutational resilience presumably maximizes evolutionary plasticity and gives the host an evolutionary advantage over viruses. It is possible that the similar mutational resilience we observed in ZED1 phenotypes gives plants an evolutionary advantage over pathogens. Furthermore, the loss of canonical kinase motifs observed in ZED1 has been hypothesized to enable evolutionary plasticity. While canonical catalytically active kinases have precise geometry for phosphotransfer, this requirement is relaxed in pseudokinases like ZED1, which may enable more freedom to evolve as an allosteric regulator [29,42]. Overall, our findings suggest that the mutational resilience of ZED1 may provide an evolutionary advantage to plants. It is likely that

mutations resulting in the loss of AtZAR1-mediated HopZ1a recognition in *N. benthamiana NbZAR1*-VIGS plants would be loss-of-function when tested in *A. thaliana*. However, we may be missing mutants that are only loss-of-function under native conditions in *A. thaliana.*

While ZAR1 orthologs are functionally conserved across angiosperms [6,7,9], our data suggests some functional divergence exists. From our library of 27 ZED1 point mutants, ZED1$^{G110D}$, ZED1$^{V170G}$, ZED1$^{T177M}$ and ZED1$^{W193R}$ were ETI-deficient with both AtZAR1 and NbZAR1, ZED1$^{E22V}$ and ZED1$^{W63R}$ were ETI-deficient with only AtZAR1, and ZED1$^{W144C}$, ZED1$^{W144R}$ and ZED1$^{R147G}$ were ETI-variable with only NbZAR1 (Figure 3, Table 2). Most surprisingly, ZED1$^{E22V}$ could not interact with HopZ1a and was ETI-sufficient with NbZAR1 but ETI-deficient with AtZAR1. Our ZED1-HopZ1a heterodimer models suggest that HopZ1a binds ZED1 through two separate interfaces, with the most likely binding interface surrounding the N-terminal α-helix (Figure 5). Multiple interaction interfaces have been demonstrated for other host-effector interactions. For example, in tomato, the NLR guarded kinase Pto interacts with two distinct domains of its recognized effector AvrPtoB [43]. In addition, we have previously demonstrated that two distinct domains of AtZAR1, AtZAR1$^{CC}$ and AtZAR1$^{LRR}$, can interact with ZED1 [7,12]. A central residue at the N-terminal binding interface is E22, which we found to be required for ZED1-HopZ1a interaction. However, despite abolishing HopZ1a interaction, ZED1$^{E22V}$ was capable triggering an HR through native NbZAR1. It is possible that our ZED1-HopZ1a interaction data only represent the N-terminal ZED1-HopZ1a interaction near the S84/T87 acetylation site, with the alternate ZED1-HopZ1a interaction near the T125/T177 acetylation site occurring transiently. If this is the case, then it is possible that the N-terminal interaction is required for AtZAR1, while NbZAR1 can be activated via the transient ZED1-HopZ1a interaction only. However, we found that the N-terminal acetylation site mutant, ZED1$^{S84A/T87A}$, is ETI-sufficient with AtZAR1 (Figure S11), suggesting that additional residues may be acetylated near the N-terminal ZED1-HopZ1a binding interface. This is possible, since previous studies identifying acetylation achieved less than 50% coverage of the ZED1 protein using LC-MS/MS [12,16]. Mutations in W144 and R147 result in ETI-variable phenotypes mediated only through NbZAR1. Mutations in both residues abolished the autoactive NbZAR1 HR phenotype (Figure 6). Both residues are near the ZED1 activation loop and could be responsible for activation loop orientation (Figure S12). These mutations may change the location of the activation loop, and this alteration could weaken the HR specifically with NbZAR1. Finally, we cannot rule out that functional differences between ZAR1 orthologs could be the result of evaluating AtZAR1 in a heterologous system, where AtZAR1 is transiently expressed and NbZAR1 has native expression.

Based on our understanding of the ZAR1-ZED1 recognition pathway, we expected that all ETI-deficient mutants could be explained by either a loss of interaction with ZAR1$^{LRR}$ or HopZ1a [7,12]. To our surprise, 2 ETI-deficient mutants, ZED1$^{T177M}$ and ZED1$^{W193R}$, could interact with AtZAR1$^{LRR}$ and with HopZ1a (Figure 2, Figure 4, Table 2). In addition, we found the ZED1$^{W193R}$ interaction with HopZ1a to be significantly stronger than the ZED1$^{WT}$-HopZ1a interaction in planta (Figure 4D). It is possible that a stronger HopZ1a interaction prevents proper activation of ZED1, causing ETI deficiency. In support of this hypothesis, we have previously shown that strengthened interaction between ZAR1 and ZED1 results in ETI-deficiency [18]. In addition, both T177 and W193 are near the ZED1 activation loop (Figure S12). This could account for our finding that ZED1$^{T177M}$ is ETI-deficient while previous work has shown that ZED1$^{T177A}$ is ETI-sufficient [7,16]. T177 is in close proximity to the ZED1 activation loop, and it is possible that ZED1$^{T177M}$ prevents proper activation of ZAR1 due to interference

with the activation loop, while ZED1$^{T177A}$ does not. While it is also possible that these mutations cause destabilization of the activation loop region, preventing proper activation of ZAR1 by ZED1, both ZED1$^{T177M}$ and ZED1$^{W193R}$ can cause autoactive HR when co-expressed with NbZAR1 (Figure 6A). Our data suggests that the NbZAR1 autoactive phenotype is caused by the ZED1 activation loop, suggesting that the activation loop is in the correct conformation in the ZED1$^{T177M}$ and ZED1$^{W193R}$ mutants.

We have previously shown that ZED1$^{T177A}$ is acetylated 85% less by HopZ1a relative to ZED1$^{WT}$ , while ZED1$^{T125A}$ is acetylated 29% less than ZED1$^{WT}$ [12]. The double acetylation mutant (ZED1$^{T125A/T177A}$) is still able to recognize HopZ1a, when both are expressed transiently in *N. benthamiana* (Baudin et al., 2017). However, ZED1 is also acetylated at residues S84 and T87 (Bastedo et al., 2019). Here, we showed that the T125/T177 acetylation site is required for AtZAR1 but not NbZAR1-mediated HR, while the S84/T87 acetylation site is not required for HopZ1a recognition through either ZAR1 ortholog (Figure S11). Combined with previous acetylation data, our finding suggests that less overall HopZ1a acetylation is required to activate NbZAR1 compared to AtZAR1. This finding provides further support that NbZAR1 is easier to activate than AtZAR1 [14,20,44].

Interspecies ZAR1-ZRK autoactive HR phenotypes prevent the expansion of ZAR1 recognition pathways outside of model species [14]. Here, we identify four separate clusters in the ZED1 protein that affect NbZAR1 autoactivity (Figure S12). These four clusters support our previous hypothesis that the ZED1 activation loop is the primary cause of NbZAR1 autoactivity, potentially through a clash with NbZAR1$^{NBD}$ [14]. Interestingly, in our AlphaFold model of ZED1 in complex with ZAR1, the ZED1 activation loop slightly displaces AtZAR1$^{NBD}$, reminiscent of the allosteric clash leading to activation of AtZAR1 by RKS1 [21,22] (Figure S7). We feel it is likely that this displacement is the most likely explanation of NbZAR1-ZED1 autoactivity. The differences in autoactivity between AtZAR1 and NbZAR1 are supported by the observation that NbZAR1 is easier to activate than AtZAR1 [20,44]. In addition, our finding that autoactivity deficient mutants cluster to sites within or close to the ZED1 activation loop further supports a ZED1-NbZAR1$^{NBD}$ clash as a mechanism for NbZAR1 autoactivity (Figure S12).

New advancements in machine learning may aid design of novel effector recognition in plants. The ZAR1 system is particularly appealing for this approach. Due to functional conservation of ZAR1 across angiosperms, guardees can be designed and tested in a heterologous system such as *N. benthamiana*, and function in a different system with a native NLR and minimal modifications. The structural data available for ZAR1, RKS1 and PBL2 enable high confidence modeling of ZRKs, and therefore ZAR1-ZRK and ZRK-effector interactions. The availability of NLR structural data has enabled rational expansion of effector recognition in other plant systems. Recently, the integrated heavy metal associated (HMA) domain of the rice NLR Pik-1 was modified to recognize new alleles of the *Magnaporthe oryzae* effector AVR-Pik resulting enhanced resistance in rice [45]. Similarly, the HMA domain of the rice NLR RGA5 was modified to recognize a new effector, AVR-PikD [46]. Expanded Pik-1 and RGA5 recognition was made possible in large part by the availability of structural data of AVR-Pik and AVR-PikD in complex with HMA domains. As machine learning based protein design, docking, and modeling tools continue to be introduced and improved, similar expansion of NLR recognition repertoires may soon be possible with limited structural data of effectors or host immunity components. In the meantime, current computational approaches combined with experimental data will aid our understanding of the molecular mechanisms of plant immunity. In a recent random mutational screen of the pepper NLR Pvr4, structural homology modeling with

AlphaFold placed mutations involved in protein stability, localization, and immune signal transduction in the context of the Pvr4 CC-domain for an expanded understanding of Pvr4 function [47]. This screen also revealed residues that are involved in NLR interaction with an ATPase signaling partner, an emerging mechanism of NLR signal transduction [48]. Here, we place random mutations that affect the ZED1-dependent NbZAR1 autoactive phenotype in the context of the ZED1 AlphaFold structure. Combining mutational and computational approaches points to the ZED1 activation loop as the probable cause of autoactivity. This further demonstrates that computational and random mutagenesis approaches can be integrated to identify novel molecular mechanisms of plant immunity.

Using our point mutant library, we identified a functional ZED1 mutant that abolishes NbZAR1 autoactivity and show for the first time that a ZAR1 recognition pathway from *A. thaliana* can function with a ZAR1 ortholog from a non-model species (Figure 6). For both NbZAR1 and SlZAR1, transient expression with ZED1 caused an autoactive HR phenotype that was abolished by the ZED1$^{N32D}$ mutation (Figure 6C). This suggests that the mechanism of interspecies ZAR1-ZRK autoactivity is conserved in the Solanaceae. Although SlZAR1 can mediate recognition of HopZ1a through ZED1$^{N32D}$, the HR phenotype observed in *N. benthamiana* was delayed relative to AtZAR1- and NbZAR1-mediated HR. Our data do not provide a clear structural explanation for this phenotype, based on SlZAR1 structure predictions. Delayed HR phenotypes have been shown to be ineffective in restricting pathogen growth [46,49,50]. It is possible that ZED1$^{N32D}$ would not confer resistance to *P. syringae* carrying HopZ1a in *S. lycopersicum.* However, ZAR1 has been previously shown to mediate effector recognition and resistance without a macroscopic HR phenotype [10]. Therefore, it is possible that SlZAR1 functions in ETI without mediating a strong HR phenotype. The functionality of SlZAR1 will need to be evaluated under native conditions. This may enable further expansion of ZAR1 immunodiversity in *A. thaliana* to non-model species.

**Methods**

*Molecular cloning*

The *ZED1* mutant library was generated with the GeneMorph II Random Mutagenesis Kit (Agilent). *ZED1* gene fragments were cloned into pDONR207 and sequenced. A Gateway-compatible version of a previously described pDex-YFP vector [7] was generated using the Gateway Conversion Kit (Invitrogen) to insert a Gateway cassette into pDex-YFP digested with *Stu*I. All *ZED1* clones with missense mutations were cloned into pDex-YFP using the Gateway LR reaction. To generate *ZED1* point mutants, *A. thaliana* codon-optimized mutations were introduced into *ZED1*-pDONR207 using the NEB Site Directed Mutagenesis kit, and confirmed by sequencing. All *ZED1* point mutants were cloned into a previously described Gateway-compatible pMac16 vector [14], which was modified from pBD to contain a C-terminal 3xFlag tag, via the LR reaction. The dexamethasone-inducible HopZ1a-HA, HopZ1a$^{C216A}$-HA, HopZ1a$^{C216A}$-cYFP, AtZAR1-5xMyc, AtZAR1$^{LRR}$-5xMyc, and NbZAR1-5xMyc constructs were previously described [7,12]. ZED1-nYFP fusions were generated by constructing a Gateway-compatible version of a previously described pDex-nYFP vector [7] using the Gateway Conversion Kit (Invitrogen) to insert a Gateway cassette into pDex-nYFP digested with *Stu*I. The ZED1 mutants tested for interaction via BiFC were cloned into Gateway-compatible pDex-nYFP via the LR reaction. *SlZAR1* was amplified from genomic DNA extracted from *Solanum lycopersicum* (Heinz) with a 5'*Xho*I site and cloned into a previously described pMac15 vector to maintain the frame for a vector-encoded 5xMyc tag [7]. The *SlZAR1$^{synth}$* construct was designed by

manually changing the codons in the first ~1000 bp of *SlZAR1* so that no stretch of 10 nucleotides was identical to *NbZAR1* and synthesized. A Gateway-compatible pMac15 vector was generated using the Gateway Conversion Kit (Invitrogen) to insert a Gateway cassette into *Stu*I digested pMac15. *SlZAR1*$^{synth}$ was inserted into Gateway-compatible pMac15 via the LR reaction. For yeast two-hybrid assays, HopZ1a$^{C216A}$ and HopF2 constructs in pEG202, and SchF2 in pJG4-5 were previously described [12]. ZED1 point mutants were amplified to add a 5′*Xho*I and a 3′*Eco*RI site, and the mutants were cloned into pJG4-5.

*N. benthamiana growth conditions, transient expression, HR assays and VIGS*

*N. benthamiana* was grown as previously described [7]. Briefly, *N. benthamiana* seeds were sterilized 10% bleach, germinated on soil, transplanted on Sunshine #1 (Sun Gro Horticulture) supplemented with 14:14:14 Osmocote, and grown at 22°C with 9 h of light.

For transient gene expression, *A. tumefaciens* GV2260 was grown overnight at 28°C in LB with 100 µg/mL rifampicin and 50 ug/mL kanamycin. Bacteria were pelleted and resuspended in induction media (10 mM MES pH 5.6 and 200 µM acetyosyringone) for 3-4 h. The optical density at 600 nm ($OD_{600}$) of each culture was adjusted to 0.25. Inoculum was infiltrated into 4-7 week *N. benthamiana* plants using a needle-less syringe. Dexamethasone (40 µM) was sprayed on leaves 24 h post infiltration. For HR assays, plants were scored 24-96 h after dexamethasone induction, depending on the experiment. HR was scored based on 4 categories as previously described (Figure 1B) [14].

Methods for *N. benthamiana* VIGS targeting *NbZAR1* have been previously described [7]. Briefly, *A. tumefaciens* was prepared as described for transient assays and infiltrated at an $OD_{600}$ of 0.25. The fourth and fifth leaves of 3-week-old *N. benthamiana* plants were completely infiltrated with bacteria using a needle-less syringe. Plants were used for transient gene expression experiments 3 weeks after initial VIGS infiltration.

*Structural model and docking analysis*

The homology model of ZED1 in complex with ZAR1 was produced using SWISS-MODEL, based on the ZAR1-RKS1 inactive structure (pdb:6j5w) [22,51]. The AlphaFold Multimer model of ZAR1 was produced using the COSMIC$^2$ webtool [27]. Models were visualized in Chimera, and contacts were assessed using a van der Waals overlap of -1 Å [52]. To generate ZED1-HopZ1a heterodimeric models, the crystal structure of HopZ1a (pdb:5klp) was used as ligand inputs and either the homology model of ZED1, or the AlphaFold model of ZED1, was used as receptor inputs in the webtools for Cluspro, Hdock, or LZerD [31–34]. Three webtools were used with defaults for unconstrained models. For constrained models, LzerD was used to generate models where the two ZED1 acetylation site pairs, S84/T87 and T125/T177, were within 7 or 9 Å of the HopZ1a catalytic site respectively [34]. These thresholds were based on recommendations by LzerD, and distances between sites specified in the constraints. To generate constrained models In Hdock, the same sites were used as for lZerD, and the default criteria were used for distance [33].

*ZED1 in vitro purification and thermal shift assay*

ZED1 protein purification was previously described (Lewis et al. 2013). In brief, a single colony of *E. coli* BL21 codon+ carrying the pET28a-6xHis:ZED1 construct was inoculated into 10 mL 2YT containing ampicillin and chloramphenicol and grown overnight at 37°C. 1/100 of the overnight culture was inoculated into 800 mL of 2YT containing ampicillin and

chloramphenicol and grown at 37°C until the $OD_{600}$ was 0.6. Expression was induced by adding 0.2 mM IPTG and incubating overnight for up to 16 h. Cells were harvested by centrifugation at 4000 g for 30 min. The pellet was resuspended in 20 mL Lysis buffer/g of cells (20 mM Tris-HCl pH 8, 500 mM NaCl, 20% glycerol, 0.1% Triton, 10 mM imidazole, 3 mM βmercaptoethanol, protease inhibitor cocktail (Roche), lysozyme). Samples were sonicated and placed on ice. The lysate was cleared by centrifugation at 35000 g for 30 min at 4°C. Next the lysate was passed through a 0.45 μM filter, and cleared supernatant was applied to 2 mL Nickel resin (Scientifix) pre-equilibated in wash buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 20% glycerol, 0.1% Triton, 10 mM imidazole, 3 mM β-mercaptoethanol, protease inhibitor cocktail (Roche)). Nickel beads were washed with 50 mL of wash buffer, followed by 20 mL of wash buffer lacking Triton. Protein was eluted with 20 mM Tris-HCl pH 8, 500 mM NaCl, 10% glycerol, 250 mM imidazole, and 3 mM β-mercaptoethanol.

Two nickel elutions were combined and loaded onto a Superdex 75 column pre-equilibrated in 20 mM Tris-HCl pH 8, 200 mM NaCl, 10% Glycerol, and 2 mM DTT. ZED1 was eluted and protein was pooled and run on a Hitrap Q column. ZED1 was eluted with a salt gradient from 30 mM to 600 mM, and the purest fractions were pooled and concentrated to 3 mg/mL. The thermal shift assay was used to examine the nucleotide binding properties of ZED1 as previously described [29,30] and 3.6 μg of ZED1 was used per reaction.

*Yeast interaction*

Methods for assaying HopZ1a-ZED1 interactions were previously described [7,12]. In brief, the RFY206 MATα strain (DupLEX-A yeast two-hybrid system; OriGene Technologies) was transformed with pEG202-$HopZ1a^{C216A}$ or pEG202-HopF2. Transformants were selected on synthetic defined (SD) Glu-His-Ura media and confirmed by colony PCR. The EGY48 MATα strain (DupLEX-A yeast two-hybrid system; OriGene Technologies) was transformed with pJG4-5-ZED1, pJG4-5-$ZED1^{E22V}$, pJG4-5-$ZED1^{W63R}$, pJG4-5-$ZED1^{T177M}$, pJG4-5-$ZED1^{W193R}$, pJG4-5-EV, or pJG4-5-SchF2, and selected on SD Glu-Trp. Yeast mating was performed in YPAD media at 30°C, and diploids were selected twice on SD Glu-Trp-His-Ura. Diploid yeast was resuspended and diluted to $OD_{600}$=10. The $5^0$ ($OD_{600}$=10), $5^{-1}$, $5^{-2}$, and $5^{-3}$ dilutions were plated on SD Glu-Leu-Trp-His-Ura and SD Gal-Leu-Trp-His-Ura + raffinose to assay for interactions.

*Western blots*

For immunoblots of transiently expressed proteins in *N. benthamiana*, 1 $cm^2$ of leaf tissue was collected 24 h after dexamethasone and homogenized in liquid nitrogen. Samples were boiled in 100 μL of 1x Laemmli buffer, run on 10% SDS-PAGE gels, and transferred to nitrocellulose membranes. Anti-Myc (Abcam ab32072), anti-GFP (Takara 632598), or anti-Flag (Sigma-Aldrich F1804) antibodies followed by appropriate secondary antibodies coupled to a horseradish peroxidase. To assay for protein expression in yeast, the EGY48 MATα strains carrying ZED1 wild type or mutant constructs were grown overnight in SD Gal+raffinose-Trp media. Protein was extracted from cultures based on the lithium acetate/sodium hydroxide extraction method [53]. Samples were boiled in 100 μL of 1x Laemmli buffer and run on 10% SDS-PAGE gels, transferred to a nitrocellulose membrane and detected using anti-HA (Roche 12CA5) antibodies followed by appropriate secondary antibodies coupled to a horseradish peroxidase.

*Confocal microscopy*

Methods for assaying HopZ1a-ZED1 interactions via BiFC were previously described [7,12]. Briefly, samples of *N. benthamiana* leaves which transiently expressed constructs were vacuum infiltrated with water and mounted on slides. Samples were imaged using a Leica SP8 confocal laser-scanning microscope with a 40x water-immersion objective (HC PL APO CS2 40x/1.10 water). Samples were excited with a 514 nm argon laser and florescence was observed using a 520-600 nm emission window. Laser power, gain, zoom, and settings were constant across experiments. Mean fluorescence intensity was quantified using the Leica Application Suite X software.

*Co-immunoprecipitation experiments*

Methods for testing AtZAR1$^{LRR}$-ZED1 interactions via co-IP were previously described [14]. Briefly, 3 cm$^2$ of tissue was collected from at least 2 separate *N. benthamiana* plants 24 h after dexamethasone induction. Tissue samples were ground using liquid nitrogen and homogenized in 1 mL IP1 buffer (50 mM HEPES, 50 mM NaCl, 10 mM ethylene diamine tetraacetic acid (EDTA), 0.2% Triton X-100 and 0.1 mg/mL Dextran [Sigma-Aldrich D1037, St. Louis, MO], pH 7.5). Samples were incubated with 5 µL Myc-Trap magnetic beads for 2.5 h on a spinning wheel at 4°C. Beads were subjected to four short washes (tubes inverted 6 times) followed by two long washes (5 min on spinning wheel) with IP2 (50 mM HEPES, 250 mM NaCl for ZED1-3xFlag interactions or 150 mM NaCl for ZED1-YFP interactions, 10 mM EDTA and 0.1% Triton X-100, pH 7.5). Both IP1 and IP2 were supplemented with a homemade protease inhibitor cocktail (1 mM PMSF, 1 mg/ml leupeptin [Sigma-Aldrich L2023], 1 mg/ml aprotinin [Sigma-Aldrich A6191], 1 mg/ml antipain [Sigma-Aldrich A1153], 1 mg/ml chymostatin [Sigma-Aldrich C7268] and 1 mg/ml pepstatin [Sigma-Aldrich P5315]). The beads and input fraction were boiled with Laemmli buffer, and each fraction was separated on a 10% SDS-PAGE gel. Proteins were detected using an anti-Myc antibody (Abcam ab32072) anti-Flag antibody (Sigma-Aldrich F1804) or anti-GFP (Takara 632598), followed by a secondary antibody coupled to horseradish peroxidase.

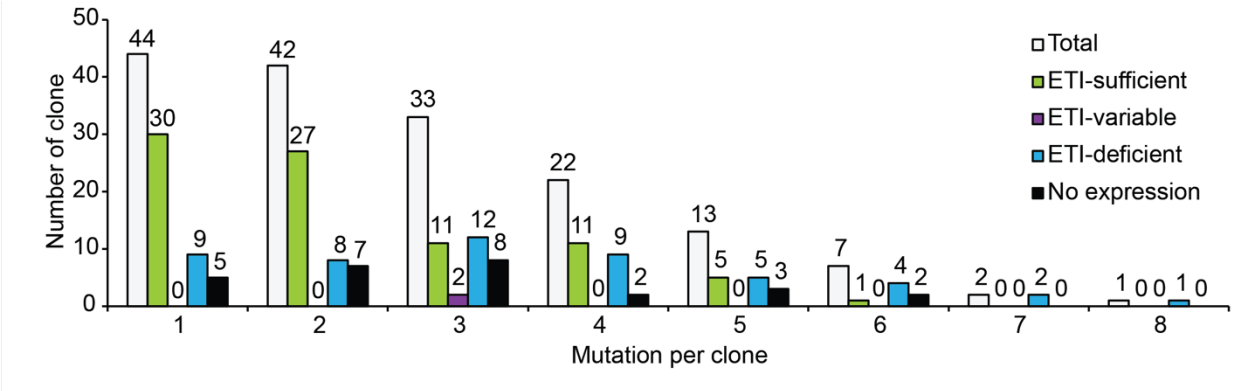
**References**


1. Yuan, M. *et al.* Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature* **592**, 105–109 (2021).
2. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
3. Grant, M. *et al.* The RPM1 plant disease resistance gene facilitates a rapid and sustained increase in cytosolic calcium that is necessary for the oxidative burst and hypersensitive cell death. *Plant J.* **23**, 441–450 (2000).
4. Jones, J. D. G., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, 1117 (2016).
5. Heath, M. C. Hypersensitive response-related death. *Plant Mol. Biol.* **44**, 321–334 (2000).
6. Lewis, J. D., Wu, R., Guttman, D. S. & Desveaux, D. Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. *PLoS Genet.* **6**, e1000894 (2010).
7. Baudin, M., Hassan, J. A., Schreiber, K. J. & Lewis, J. D. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol.* **174**, 2038–2053 (2017).
8. Schultink, A., Qi, T., Bally, J. & Staskawicz, B. Using forward genetics in Nicotiana

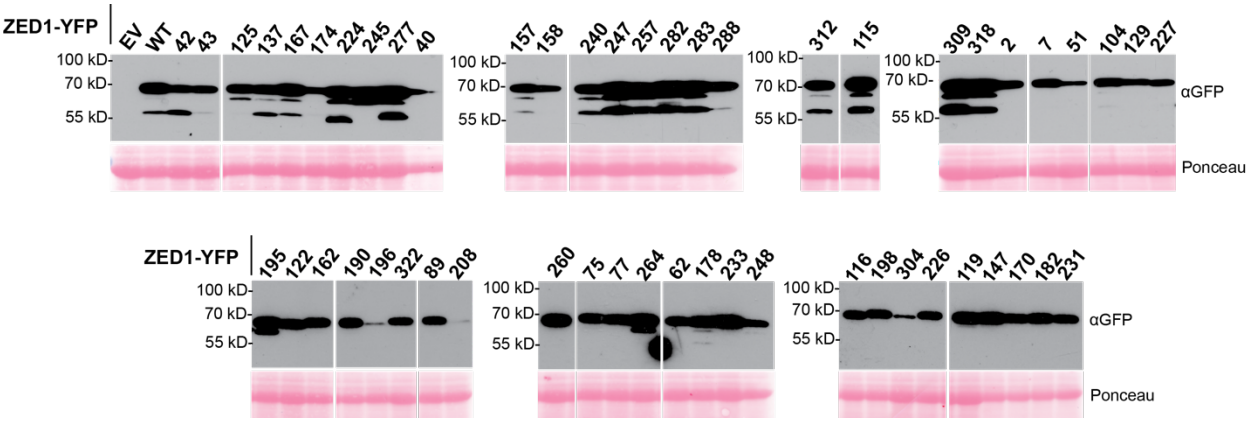
benthamiana to uncover the immune signaling pathway mediating recognition of the Xanthomonas perforans effector XopJ4. *New Phytol.* **221**, 1001–1009 (2019).
9. Gong, Z. *et al.* The origin and evolution of a plant resistosome. *Plant Cell* **34**, 1600–1620 (2022).
10. Laflamme, B. *et al.* The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. *Science* **367**, 763–768 (2020).
11. Seto, D. *et al.* Expanded type III effector recognition by the ZAR1 NLR protein using ZED1-related kinases. *Nat. Plants* **3**, 25–28 (2017).
12. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the Pseudomonas syringae type III effector HopZ1a. *Proc. Natl. Acad. Sci. USA* **110**, 18722–18727 (2013).
13. Wang, G. *et al.* The decoy substrate of a pathogen effector and a pseudokinase specify pathogen-induced modified-self recognition and immunity in plants. *Cell Host Microbe* **18**, 285–295 (2015).
14. Diplock, N. *et al.* Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1. *Plant. Cell Environ.* **46**, 2238–2254 (2023).
15. Seto, D. *et al.* The small molecule Zaractin activates ZAR1-mediated immunity in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **118**, 1–8 (2021).
16. Bastedo, D. P. *et al.* Perturbations of the ZED1 pseudokinase activate plant immunity. *PLoS Pathog.* **15**, e1007900 (2019).
17. Liu, C. *et al.* Two Arabidopsis receptor-like cytoplasmic kinases SZE1 and SZE2 associate with the ZAR1–ZED1 complex and are required for effector-triggered immunity. *Mol. Plant* **12**, 967–983 (2019).
18. Baudin, M. *et al.* A natural diversity screen in Arabidopsis thaliana reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex. *Plant. Cell Environ.* **44**, 629–644 (2021).
19. Wang, Z. *et al.* Arabidopsis ZED1-related kinases mediate the temperature-sensitive intersection of immune response and growth homeostasis. *New Phytol.* **215**, 711–724 (2017).
20. Baudin, M., Schreiber, K. J., Martin, E. C., Petrescu, A. J. & Lewis, J. D. Structure–function analysis of ZAR1 immune receptor reveals key molecular interactions for activity. *Plant J.* **101**, 352–370 (2019).
21. Wang, J. *et al.* Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science* **364**, eaav5870 (2019).
22. Wang, J. *et al.* Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science* **364**, eaav5868 (2019).
23. Zheng, X. *et al.* Two plant NLR proteins confer strain-specific resistance conditioned by an effector from Pseudomonas syringae pv. actinidiae. *J. Genet. Genomics* **49**, 823–832 (2022).
24. Adachi, H. *et al.* Jurassic NLR: Conserved and dynamic evolutionary features of the atypically ancient immune receptor ZAR1. *Plant Cell* 1–50 (2023) doi:10.1093/plcell/koad175.
25. Mccullum, E. O., Williams, B. a R., Zhang, J. & Chaput, J. C. Error based PCR mutagenesis protocols. *Methods Mol. Biol.* **634**, 103–109 (2010).
26. Shan, L. *et al.* The HopPtoF locus of Pseudomonas syringae pv. tomato DC3000 encodes a type III chaperone and a cognate effector. *Mol. Plant-Microbe Interact.* **17**, 447–455

(2004).
27. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021.10.04.463034 (2022) doi:10.1101/2021.10.04.463034.
28. Hu, M., Qi, J., Bi, G. & Zhou, J. M. Bacterial Effectors Induce Oligomerization of Immune Receptor ZAR1 In Vivo. *Mol. Plant* **13**, 793–801 (2020).
29. Murphy, J. M. *et al.* A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. *Biochem. J.* **457**, 323–334 (2014).
30. Lucet, I. S. *et al.* Determination of pseudokinase-ligand interaction by a fluorescence-based thermal shift assay. **4**, 1–23 (2014).
31. Zhang, Z. M. *et al.* Structure of a pathogen effector reveals the enzymatic mechanism of a novel acetyltransferase family. *Nat. Struct. Mol. Biol.* **23**, 847–852 (2016).
32. Kozakov, D. *et al.* The ClusPro web server for protein–protein docking. *Nat. Protoc.* **12**, 255–278 (2017).
33. Yan, Y., Tao, H., He, J. & Huang, S. Y. The HDOCK server for integrated protein–protein docking. *Nat. Protoc.* **15**, 1829–1852 (2020).
34. Christoffer, C. *et al.* LZerD webserver for pairwise and multiple protein-protein docking. *Nucleic Acids Res.* **49**, W359–W365 (2021).
35. Zhang, N. *et al.* MutaBind2: predicting the impacts of single and multiple mutations on protein-protein interactions. *iScience* **23**, 100939 (2020).
36. Van de Weyer, A. L. *et al.* A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis thaliana. *Cell* **178**, 1260-1272.e14 (2019).
37. Dillon, M. M. *et al.* Molecular evolution of Pseudomonas syringae type III secreted effector proteins. *Front. Plant Sci.* **10**, 1–18 (2019).
38. Martel, A. *et al.* The ETS-ETI cycle: evolutionary processes and metapopulation dynamics driving the diversification of pathogen effectors and host immune factors. *Curr. Opin. Plant Biol.* **62**, 102011 (2021).
39. Karasov, T. L. *et al.* The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* **512**, 436–440 (2014).
40. Tenthorey, J. L., Young, C., Sodeinde, A., Emerman, M. & Malik, H. S. Mutational resilience of antiviral restriction favors primate TRIM5α in host-virus evolutionary arms races. *Elife* **9**, 1–36 (2020).
41. Tenthorey, J. L., Emerman, M. & Malik, H. S. Evolutionary landscapes of host-virus arms races. *Annu. Rev. Immunol.* **40**, 271–294 (2022).
42. Murphy, J. M. *et al.* Insights into the evolution of divergent nucleotide-binding mechanisms among pseudokinases revealed by crystal structures of human and mouse MLKL. *Biochem. J.* **457**, 369–377 (2014).
43. Mathieu, J., Schwizer, S. & Martin, G. B. Pto kinase binds two domains of AvrPtoB and its proximity to the effector E3 ligase determines if it evades degradation and activates plant immunity. *PLoS Pathog.* **10**, 24–27 (2014).
44. Harant, A., Pai, H., Sakai, T., Kamoun, S. & Adachi, H. A vector system for fast-forward studies of the HOPZ-ACTIVATED RESISTANCE1 (ZAR1) resistosome in the model plant Nicotiana benthamiana. *Plant Physiol.* **188**, 70–80 (2022).
45. Maidment, J. H. R. *et al.* Effector target-guided engineering of an integrated domain expands the disease resistance profile of a rice NLR immune receptor. *Elife* **12**, 1–25 (2023).
46. Césari, S. *et al.* New recognition specificity in a plant immune receptor by molecular

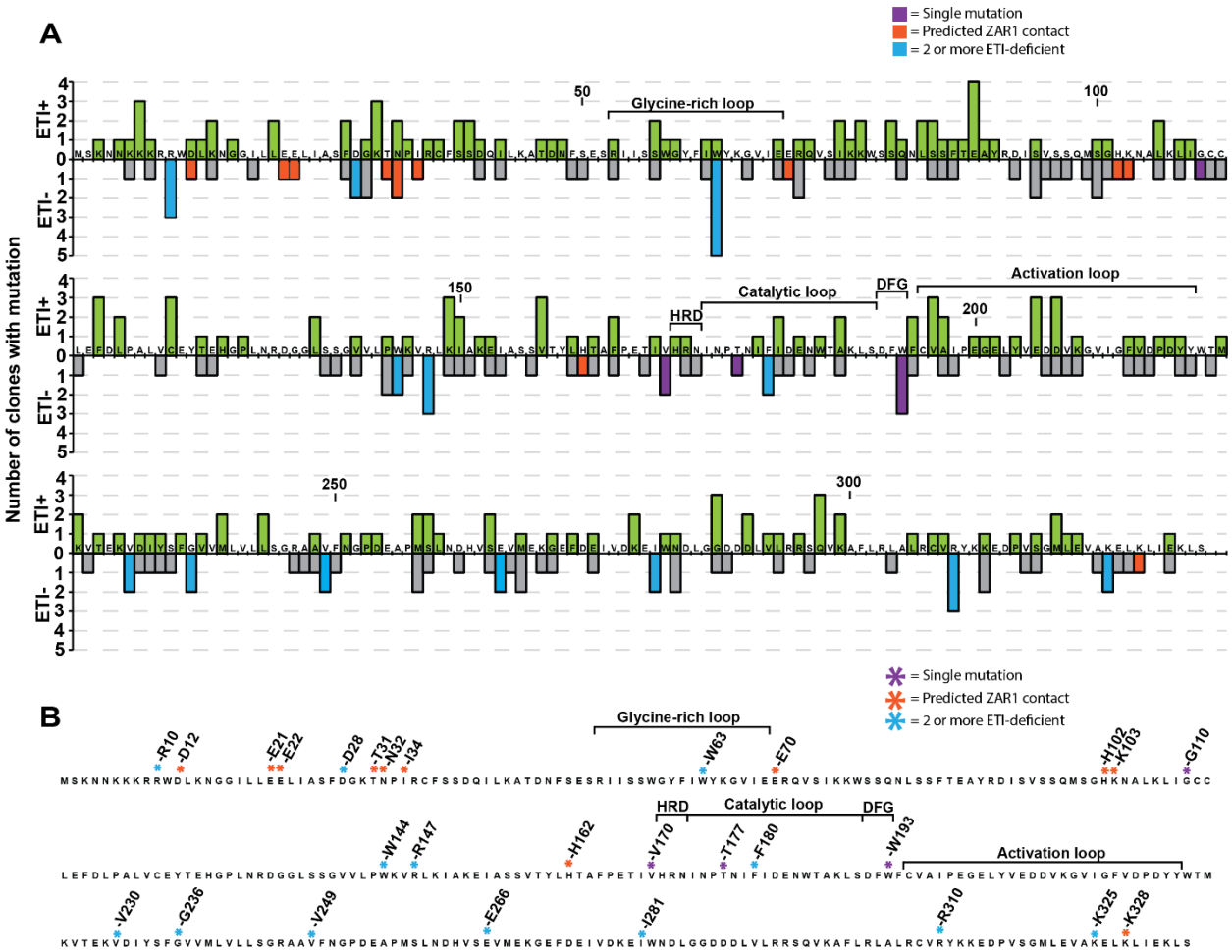
engineering of its integrated domain. *Nat. Commun.* **13**, 1–13 (2022).
47. Lee, H. Y. *et al.* Protein stability governed by α1–2 helices in Pvr4 is essential for localization and cell death. *Plant J.* 1–14 (2023) doi:10.1111/tpj.16388.
48. Lee, H. Y. *et al.* Plasma membrane-localized plant immune receptor targets H+-ATPase for membrane depolarization to regulate cell death. *New Phytol.* **233**, 934–947 (2022).
49. Pottinger, S. E. *et al.* Optimizing the PBS1 decoy system to confer resistance to potyvirus infection in Arabidopsis and soybean. *Mol. Plant-Microbe Interact.* **33**, 932–944 (2020).
50. Helm, M. *et al.* Engineering a decoy substrate in soybean to enable recognition of the soybean mosaic virus NIa protease. *Mol. Plant-Microbe Interact.* **32**, 760–769 (2019).
51. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
52. Pettersen, E. F. *et al.* UCSF Chimera-A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
53. Zhang, T. *et al.* An improved method for whole protein extraction from yeast *Saccharomyces cerevisiae*. *Yeast* **28**, 795–798 (2011).



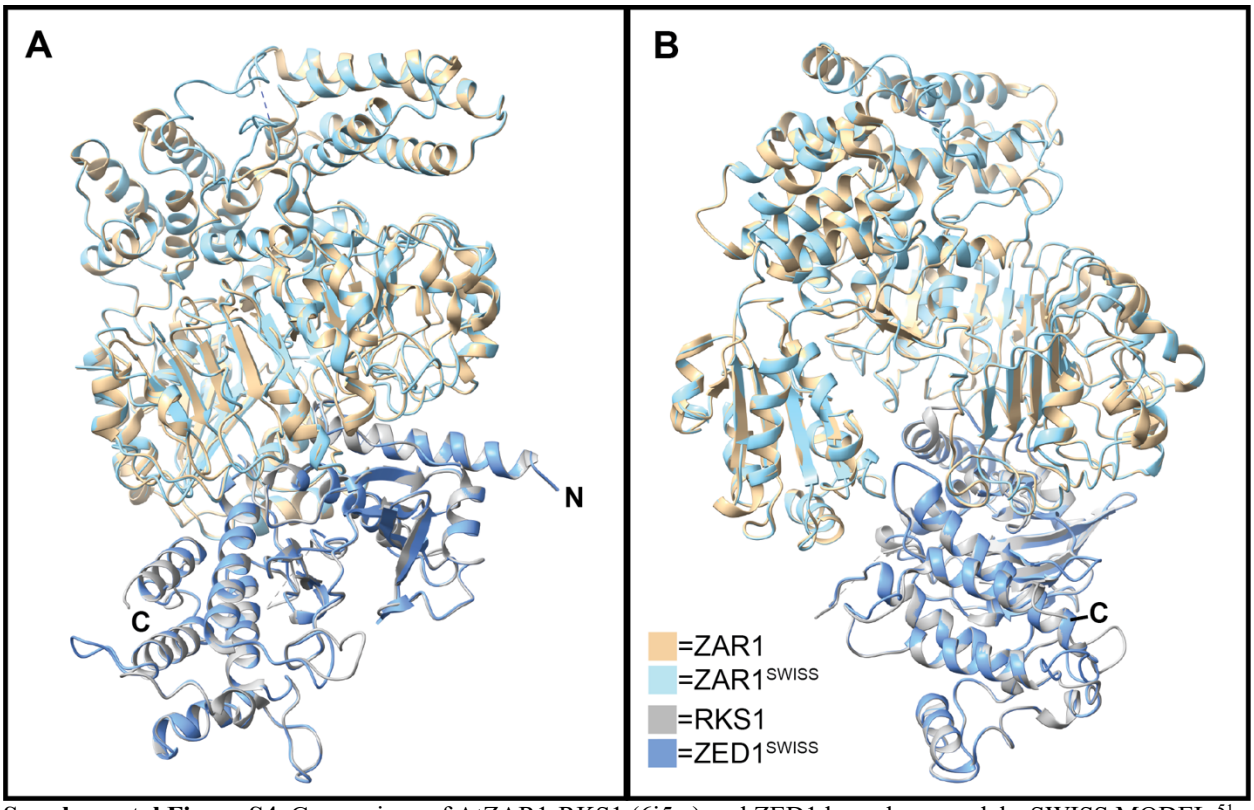

**Supplemental Figure S1.** Summary of phenotypes associated with ZED1 error prone PCR library. Each ZED1 mutant was co-expressed with HopZ1a in *N. benthamiana*, and the HR was scored on a 4-point scale (see Figure 1B) 24 hours after dexamethasone induction (n=5-6 plants, see Supplemental Table S1). For all ETI-deficient mutants, ZED1 expression was tested by western blot (see Figure S2).



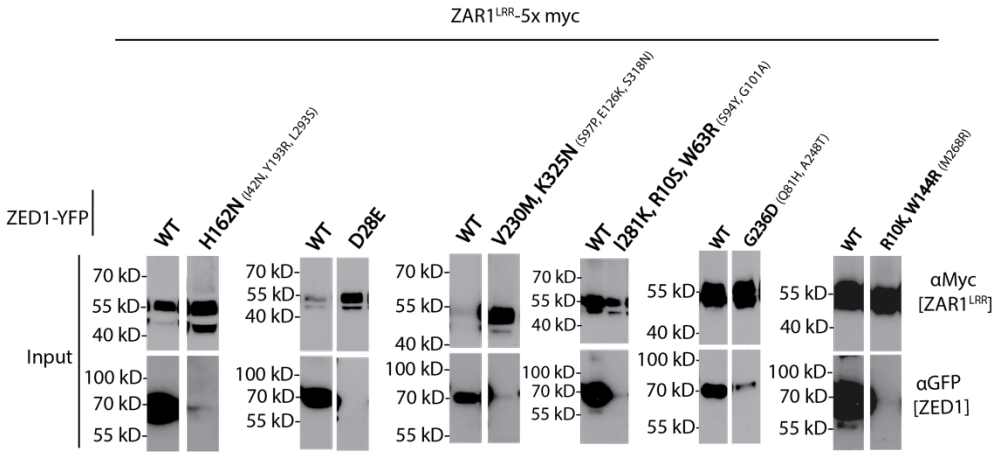

**Supplemental Figure S2.** Western blots showing expression of ETI-deficient ZED1 mutants. ZED1-YFP mutants were transiently expressed in *N. benthamiana,* and protein was detected via Anti-GFP western blots. ZED1 mutants with weak or strong protein expression are shown. Predicted molecular mass of ZED1-YFP is 67 kDa.



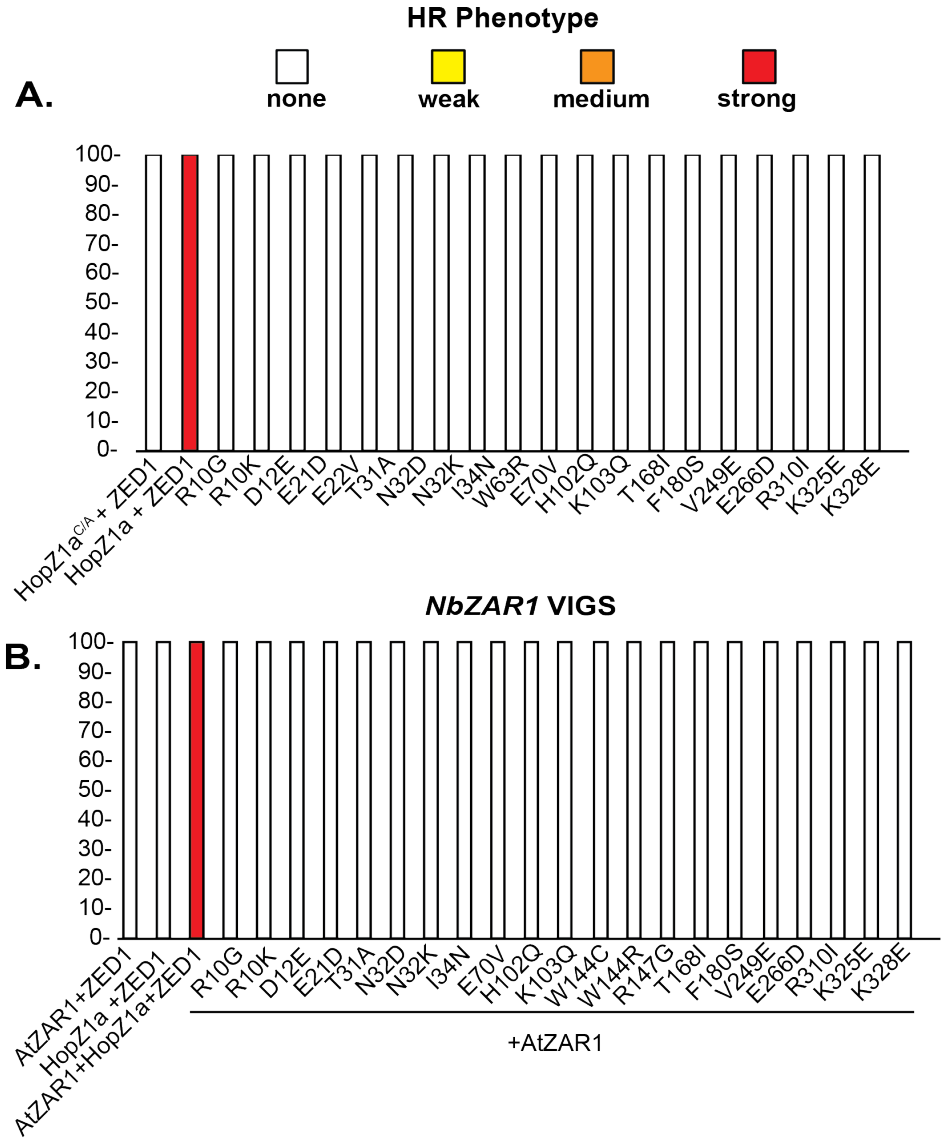

**Supplemental Figure S3.** Location of error prone library candidate mutations in ZED1 protein. (A) Full-length ZED1 clones with non-synonymous mutations are plotted on the corresponding residue in ZED1 based on ETI phenotypes (see Supplemental Table S1, Figure 1). Bars are colored by criteria used to prioritize clones and residues for further analysis. (B) Location of candidate mutations on the ZED1 protein is indicated by asterisks with colors representing the selection criteria as in part A.



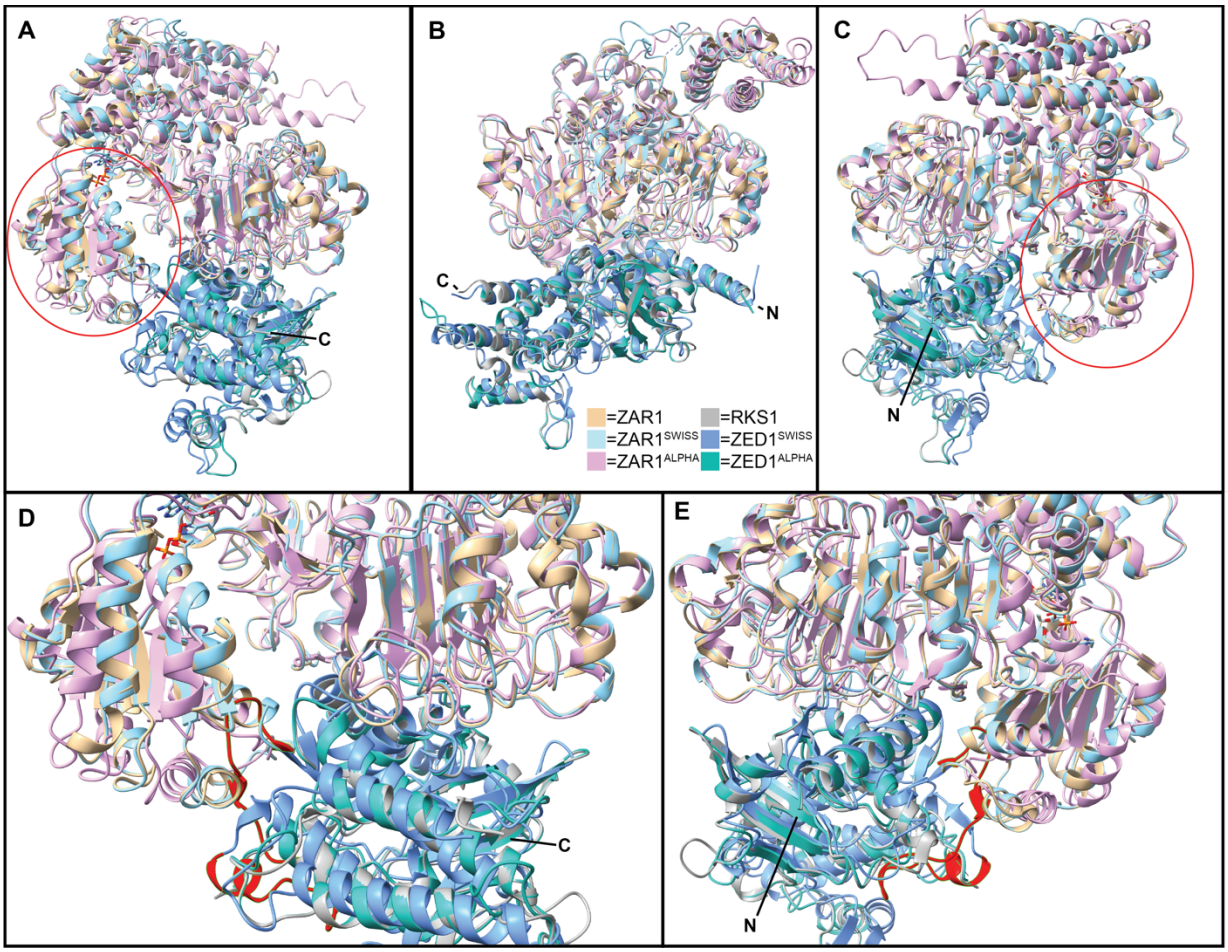

**Supplemental Figure S4.** Comparison of AtZAR1-RKS1 (6j5w) and ZED1 homology models. SWISS MODEL [51] was used to generate the ZED1$^{SWISS}$ homology model of ZAR1 (light blue), and ZED1 (dark blue), based on the inactive structure (pdb: 6j5w) of ZAR1 (beige) in complex with RKS1 (grey).



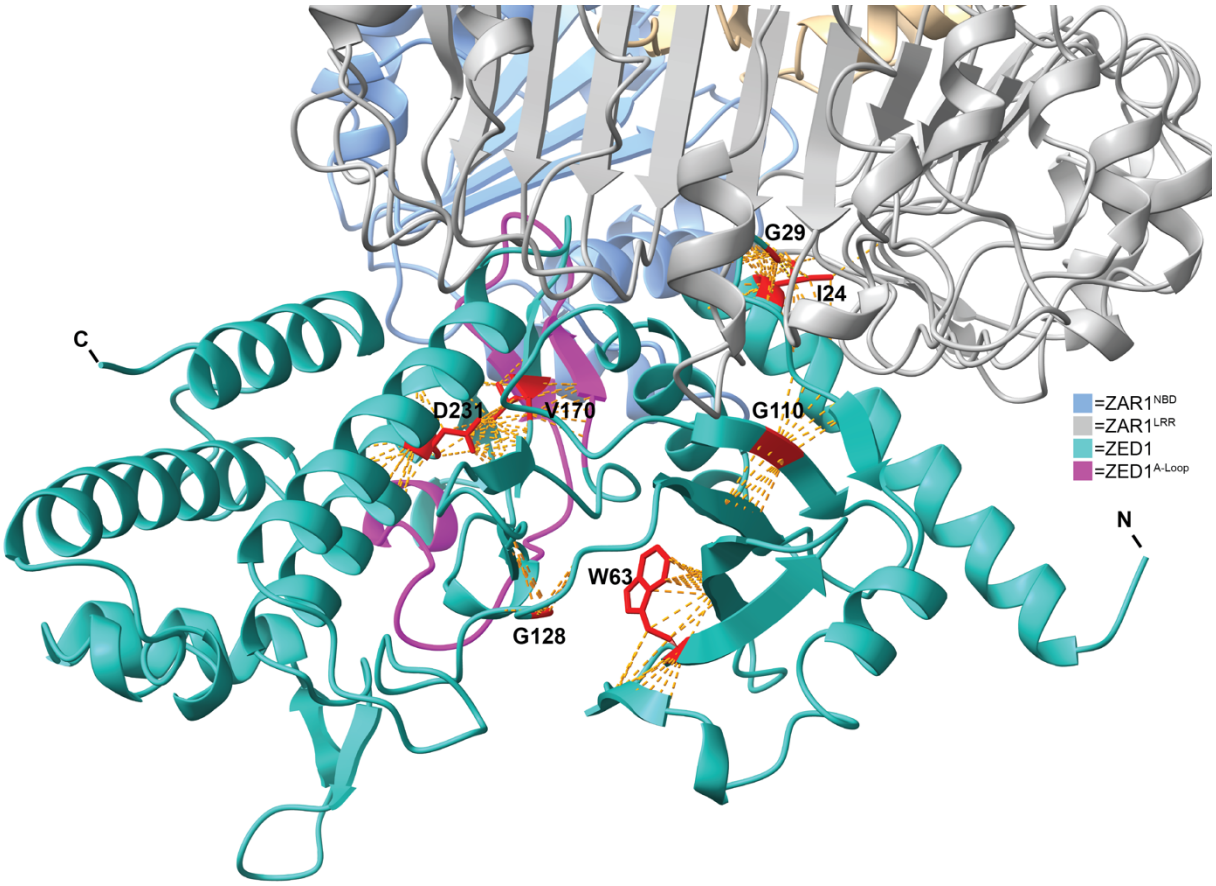

**Supplemental Figure S5.** Western blots of ZED1 mutants excluded from further analysis due to lack of protein accumulation. Each ZED1 mutant prioritized based on ETI phenotypes was transiently co-expressed with AtZAR1$^{LRR}$ in *N. benthamiana.* Anti-Myc western blots indicate that protein is present in all samples, while anti-GFP western blots show mutant protein accumulation relative to ZED1$^{WT}$.



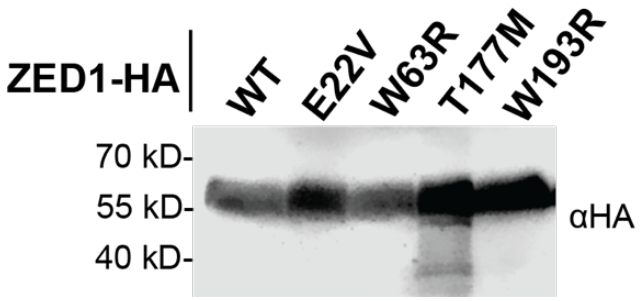

**Supplemental Figure S6.** ZED1 point mutants do not cause an autoactive HR phenotype when transiently expressed in *N. benthamiana* or when co-expressed with AtZAR1 in *N. benthamiana NbZAR1* VIGS plants. ETI-sufficient ZED1 mutants were (a) transiently expressed in *N. benthamiana* or (b) transiently co-expressed with AtZAR1 in *N. benthamiana NbZAR1*-VIGS plants [7]. The HR was scored on a 4-point scale (Figure 1B) 24 hours after dexamethasone induction (n=13-17 plants). To ensure VIGS of *NbZAR1* was successful, HopZ1a and ZED1 were co-expressed in each *NbZAR1*-VIGS plant [7]. Data are compiled from at least three separate experiments.



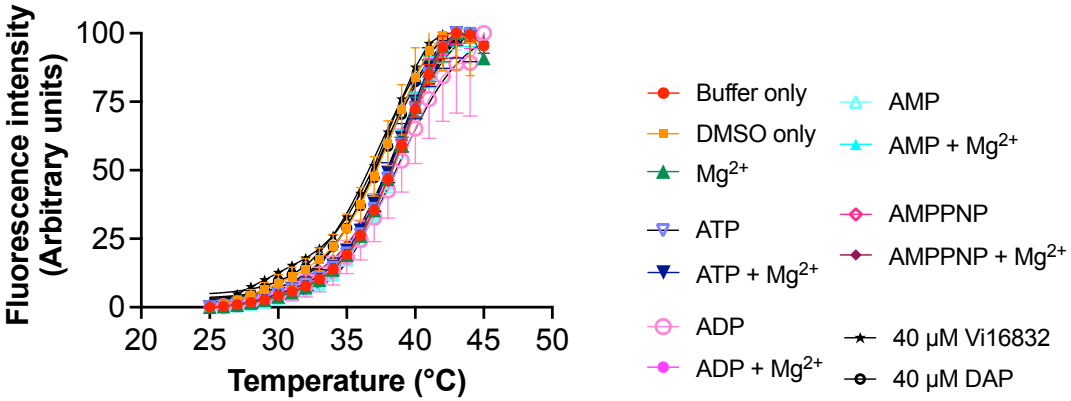

**Supplemental Figure S7.** Comparison of ZAR1-RKS1 (6j5w), ZED1 homology, and ZED1 AlphaFold protein models. The ZED1$^{SWISS}$ homology model of ZAR1 (light blue), and ZED1 (dark blue), was generated in SWISS MODEL [51], based on the inactive structure (pdb: 6j5w) of ZAR1 (beige) in complex with RKS1 (grey). The AlphaFold model of ZAR1 (pink) in complex with ZED1 (aqua-blue) was generated using AlphaFold multimer [27]. The activation loop of ZED1$^{ALPHA}$ is shown in red. Red circles indicate the location of ZAR1$^{NBD}$. Proteins were visualized in Chimera [52].




**Supplemental Figure S8.** Residues important for ZAR1-ZED1 interaction displayed on AlphaFold model of ZED1. The AlphaFold model of ZAR1 in complex with ZED1 (aqua-blue) was generated using AlphaFold multimer [27]. The ZAR1$^{LRR}$ and ZAR1$^{NBD}$ are shown in grey and light blue respectively. The ZED1 activation loop is shown in magenta. Residues required for ZAR1 interaction, and their contacts are shown in red and orange respectively. Proteins were visualized in Chimera [52].




**Supplemental Figure S9.** Western blot showing expression of ZED1 single mutants in yeast. The EGY48 MATα strains carrying ZED1, or mutant constructs was grown overnight in SD inducing media. Protein was extracted from cultures based on the lithium acetate/sodium hydroxide extraction method [53]. Protein was detected via anti-HA western blot.




**Supplemental Figure S10.** ZED1 does not bind nucleotides or cations in thermal shift assay. ZED1 protein was expressed in *E. coli* and purified. The Thermal Shift Assay was used to examine the nucleotide binding properties of ZED1 as previously described [29,30].

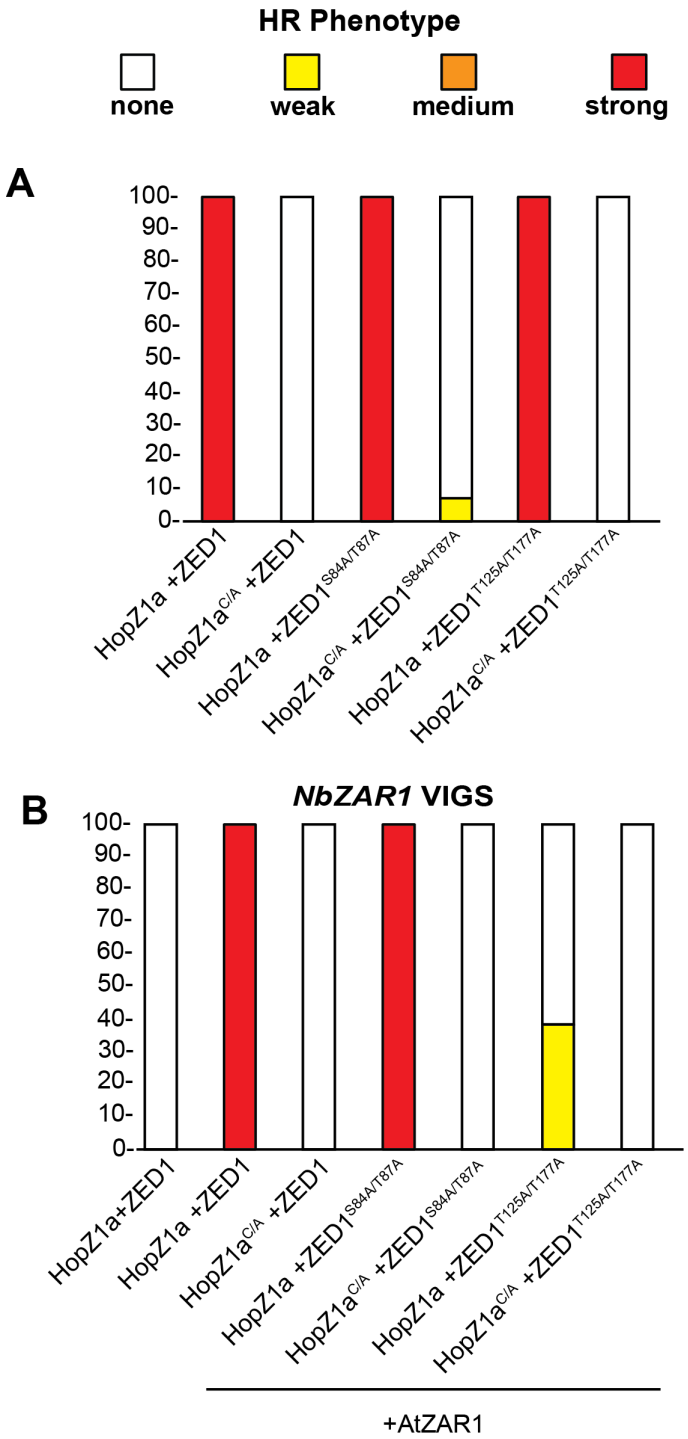
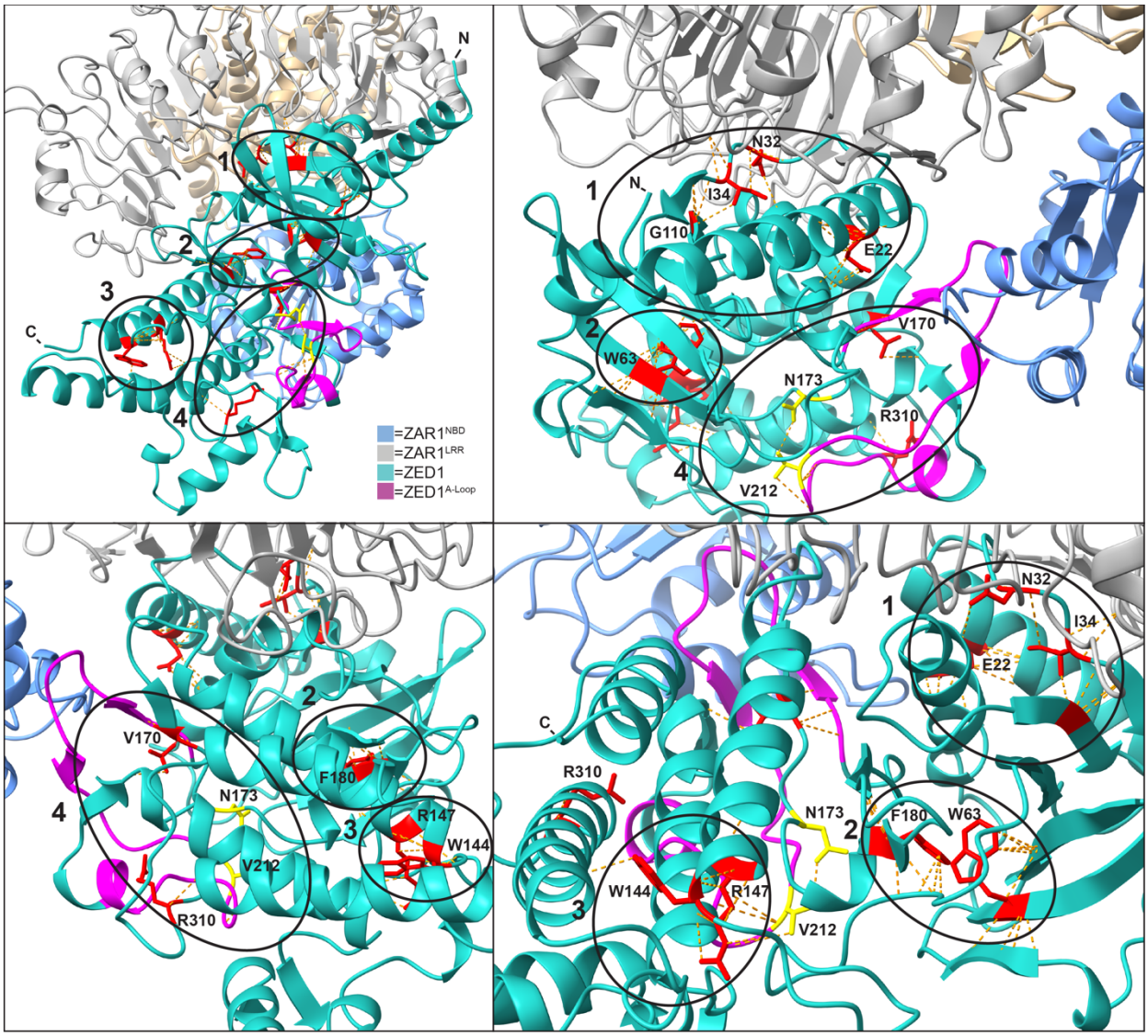
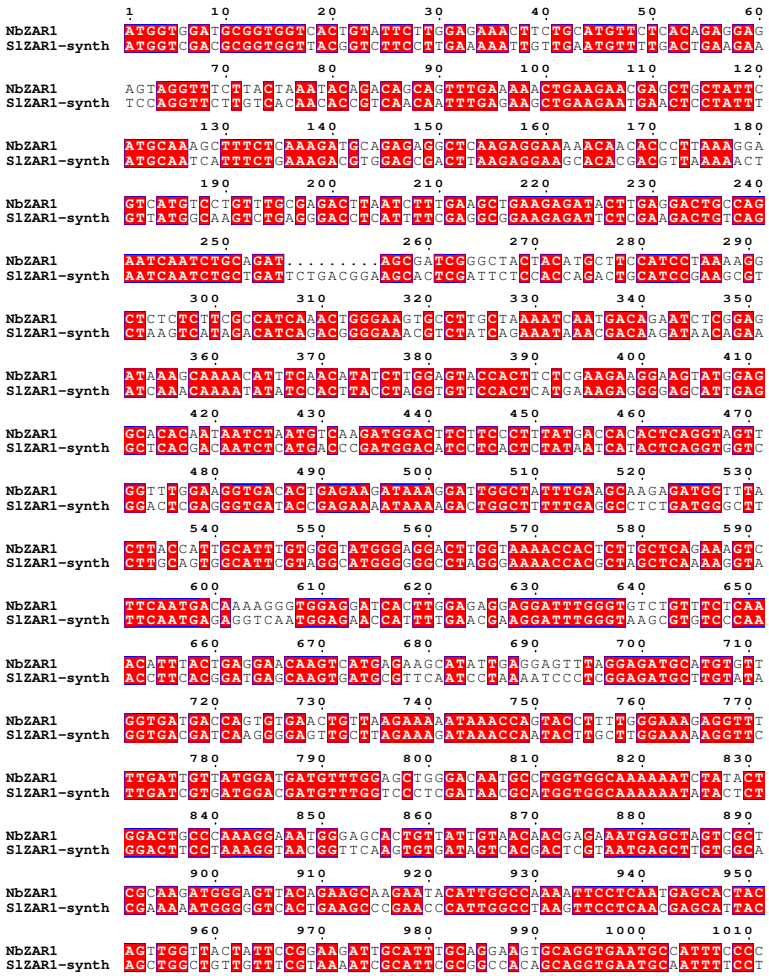


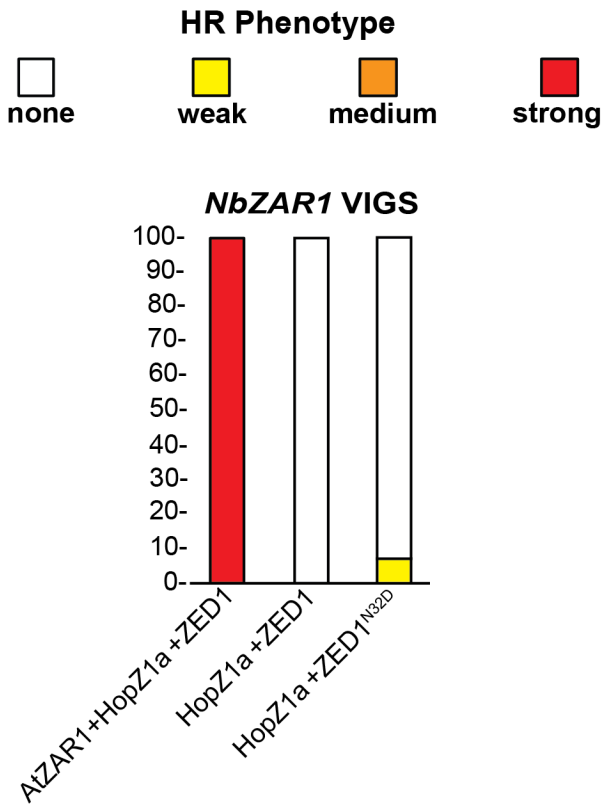

**Figure S11.** ZED1 acetylation sites are required for AtZAR1 but not NbZAR1-mediated recognition of HopZ1a. ZED1 acetylation site mutants were coexpressed with either HopZ1a or HopZ1a$^{C216A}$ in (A) *N. benthamiana* or (B) co-expressed with AtZAR1 in *N. benthamiana NbZAR1*-VIGS plants [7]. The HR was scored on a 4-point scale (Figure 1B) 24 hours after dexamethasone induction (n=13-14 plants). To ensure VIGS of *NbZAR1* was successful, HopZ1a and ZED1 were co-expressed in each *NbZAR1*-VIGS plant [7]. Data are compiled from three separate experiments.




**Supplemental Figure S12.** Mutations leading to loss of NbZAR1 autoactive HR cluster around ZED1 activation loop. Four separate views of the AlphaFold model of ZED1 (aqua-blue) in complex with ZAR1$^{LRR}$ (grey) and ZAR1$^{NBD}$ (blue). The ZED1 activation loop is shown in purple. Mutations affecting ZED1-NbZAR1 autoactivity are shown in red with contacts of these residues shown in orange. Yellow residues have been previously shown to result in AtZAR1 autoactivity when mutated [16,19]. Proteins were visualized in Chimera [52].




**Supplemental Figure S13.** Nucleotide alignment of NbZAR1 and SlZAR1$^{synth}$ in the region targeted by *NbZAR1* VIGS. Codons in SlZAR1$^{synth}$ were manually switched to create synonymous mutations so that the transcript is not silenced when transiently expressed in *NbZAR1*-VIGS plants.




**Supplemental Figure S14.** ZED1$^{N32D}$ does not cause HR through another NLR in *N. benthamiana*. ZED1$^{N32D}$ was coexpressed with HopZ1a in *NbZAR1*-VIGS plants [7]. The HR was scored on a 4-point scale (Figure 1B) 96 hours after dexamethasone induction (n= plants). To ensure VIGS of *NbZAR1* was successful, HopZ1a and ZED1 were co-expressed in each *NbZAR1*-VIGS plant [7]. ZED1 was coexpressed with HopZ1a$^{C216A}$ as a negative control for the HR. Data are compiled from three separate experiments.

## CONCLUSION

Here, we use diversity in ZED1 to identify determinants for ZAR1 recognition of HopZ1a, and ZAR1 autoactivation. We identified a mutation in ZED1 that enables recognition of HopZ1a through tomato ZAR1, a key step towards rational design of ZAR1 recognition pathways for *P. syringae* resistance.

## CHAPTER V: *Solanum pimpinellifolium* exhibits complex genetic resistance to *Pseudomonas syringae pv. tomato*

### PREFACE

This work was conducted in collaboration with Jana A. Hassan, Ilea J. Chau-Ly, Jamie Calma, Elizabeth Boville, Steven Yee, Taylor M. Harris, and Jennifer D. Lewis. I performed the bioinformatic analyses with Jana A. Hassan. Jana A. Hassan performed most experiments, Jamie Calma, Elizabeth Boville, Steven Yee, Taylor M. Harris contributed to seedling flood assays, and Taylor M. Harris contributed to HR assays. I analyzed the data along with Jana A. Hassan, Ilea J. Chau-Ly and Jennifer D. Lewis, and wrote the manuscript with Jana A. Hassan, Ilea J. Chau-Ly and Jennifer D. Lewis, with assistance from Jamie Calma, Elizabeth Boville, Steven Yee, Taylor M. Harris. The data presented here has been reproduced with permission of all collaborators.

This work uses natural diversity in the *Solanum* genus to identify resistance to *P. syringae*. This is a key step towards rational design of *P. syringae* resistance in crops.


### Abstract

*Pseudomonas syringae pv. tomato* (*Pst*) is the causal agent of bacterial speck disease in tomatoes. The *Pto/Prf* gene cluster from *Solanum pimpinellifolium* was introgressed into many modern tomato cultivars and historically provided protection against *Pst* race 0 strains. However, virulent *Pst* race 1 strains that evade *Pto*-mediated immunity now predominate in tomato-growing regions worldwide. Here we report the identification of resistance to a *Pst* race 1 strain (*Pst*19) in the wild tomato accession *S. pimpinellifolium* LA1589 (hereafter LA1589), using our rapid high-throughput seedling screen. LA1589 supports less bacterial growth than cultivars, and does not exhibit a hypersensitive response to *Pst*19. We tested an existing set of 87 Inbred Backcross Lines (IBLs) derived from a cross between *Solanum esculentum* E-6203 x *Solanum pimpinellifolium* LA1589 for resistance to *Pst*19. Using single-marker analysis, we identified five genomic regions associated with resistance. Bacterial growth assays on IBLs confirmed that these regions contribute to resistance *in planta*. We also mapped candidate genes associated with resistance in a cross between the *Solanum lycopersicum* var. *lycopersicum* cultivar Heinz BG-1706 and *S. pimpinellifolium* LA1589. By comparing candidates from the two mapping approaches, we were able to identify 5 QTL and 11 candidate genes in LA1589 for a role in resistance to *Pst*19. This work assists molecular marker-assisted breeding for protection of plants from disease.


### Introduction

Bacterial speck disease of tomato is a persistent problem for tomatoes grown throughout the world, affecting both the marketability and yield of fresh-market and processing tomatoes (Yunis et al. 1980; Kunkeaw et al. 2010). *Pseudomonas syringae pv. tomato* (*Pst*) is the causal agent of bacterial speck on tomato. It is primarily a foliar pathogen that grows as an epiphyte on the plant surface, and gains entry through the stomata to proliferate in the plant apoplast (Hirano and Upper 2000; Xin et al. 2018). Plants defend themselves against pathogens through pattern-triggered immunity (PTI), where conserved microbial molecules are recognized, and effector-triggered immunity (ETI), where specific effectors are recognized. PTI occurs at the surface of the plant cell, where pattern recognition receptor proteins (PRRs) recognize highly conserved pathogen-associated molecular patterns (Couto and Zipfel 2016). As a result of PTI, the host

upregulates a variety of defense responses, including reinforcement of the cell wall, production of reactive oxygen species, and stomatal closure (Couto and Zipfel 2016; DeFalco and Zipfel 2021). To overcome PTI, bacteria utilize a type III secretion system to translocate effector proteins directly into the plant cell, where they target components of PTI and promote bacterial virulence (Lewis et al. 2009; Block and Alfano 2011; Schreiber et al. 2021). ETI occurs upon recognition of effectors by nucleotide-binding-site leucine-rich repeat receptor (NLRs, also called Resistance (R) proteins), which triggers a rapid and robust secondary immune response (DeYoung and Innes 2006; Jones et al. 2016; Schreiber et al. 2016). This is often associated with localized cell death at the site of infection, termed the hypersensitive response (HR) (Heath 2000). In addition to the gene-for-gene resistance often associated with ETI, plants can also exhibit quantitative disease resistance (QDR), in which multiple loci partially contribute to disease resistance (Poland et al. 2009; Boyd et al. 2013; French et al. 2016). QDR typically results in less disease, rather than an absence of disease, and can be affected by environmental conditions (French et al. 2016). It is considered to confer stable resistance in the field, and is therefore very useful in breeding programs (Poland et al. 2009).

Historically, bacterial speck on tomato was primarily caused by *Pst* race 0 strains, which translocate a suite of type III secreted effector proteins (T3SEs) that includes AvrPto and AvrPtoB (Pedley and Martin 2003). In the 1930s, the *Pto/Prf* gene cluster was introgressed from the wild tomato species *Solanum pimpinellifolium* into processing tomato cultivars to provide protection against bacterial speck disease caused by *Pst* race 0 strains (Pilowsky and Zutra 1982; Kerr and Cook 1983; Pitblado and Macneill 1983; Pedley and Martin 2003). Pto, a protein kinase, recognizes AvrPto and AvrPtoB through direct interaction, and works together with Prf, a NLR, to activate ETI (Ronald et al. 1992; Martin et al. 1993; Salmeron et al. 1996; Tang et al. 1996). However, the *Pto/Prf* gene cluster does not recognize *Pst* race 1 strains, which have emerged as the prevalent strains throughout the world (Kunkeaw et al. 2010; Cai et al. 2011). *Pst* race 1 strains are differentiated from *Pst* race 0 strains by the loss, mutation, or post-transcriptional down-regulation of AvrPto and AvrPtoB, which abrogates recognition by Pto/Prf (Lin et al. 2006; Almeida et al. 2009; Kunkeaw et al. 2010).

Wild relatives of crop plants are excellent sources of natural genetic diversity for traits of interest, including pathogen resistance (Zamir 2001). Several studies in adult plants have identified resistance to *Pst* race 1 strains in wild relatives of tomato. The wild tomato accession *S. habrochaites* LA1777 demonstrates resistance to *Pst*A9, a *Pst* race 1 isolate found in California, and four quantitative trait loci (QTL) associated with resistance were identified using a population of introgression lines (Thapa et al. 2015). Another *S. habrochaites* accession, LA2109, demonstrates resistance to the *Pst* race 1 strain T1 (*Pst*T1) (Bao et al. 2015). Two QTL and a candidate gene for resistance were identified using a mapping-by-sequencing approach in LA2109. Recently, researchers identified the NLR Ptr1, which is able to detect the activity of the effector AvrRpt2 and trigger resistance to *Pst*T1 in *S. lycopersicoides* LA2951 (Mazo-Molina et al. 2019; Mazo-Molina et al. 2020).

To rapidly screen for resistance to *P. syringae* in wild tomato accessions, we developed a high-throughput seedling-based flood assay which faithfully recapitulates adult phenotypes (Hassan et al. 2017; Hassan et al. 2020). We demonstrated that seedlings of cultivars containing the *Pto/Prf* cluster recognize *Pst*DC3000, resulting in seedling survival, reduced bacterial growth, a hypersensitive response, and a rapid increase in ion leakage (Hassan et al. 2017). Seedlings of cultivars that lack the *Pto/Prf* cluster are susceptible to *Pst*DC3000, and support high levels of bacterial growth with eventual death of the seedlings (Hassan et al. 2017). Using

this screen, we identified two additional wild tomato accessions which demonstrate strong resistance to *Pst*19, *S. neorickii* LA1329 and *S. habrochaites* LA1253 (Hassan et al. 2017). *Pst*19 is a hypervirulent strain of *P. syringae* closely related to *Pst*T1 (Kunkeaw et al. 2010). *S. neorickii* LA1329 displays genetically complex resistance to *Pst*19 in both seedlings and adult plants.

In this study, we identified a wild tomato line, *S. pimpinellifolium* LA1589, which exhibits resistance to *Pst*19 in both seedlings and adults. Although many *S. pimpinellifolium* accessions possess the *Pto*/*Prf* gene cluster (Pitblado and Macneill 1983; Pedley and Martin 2003), we found that resistance to *Pst*19 was not dependent on Pto/Prf. Additionally, we found that LA1589 does not exhibit an HR or levels of bacterial growth that are characteristic of ETI, suggesting it has QDR to *Pst*19. We used a pre-existing population of *S. pimpinellifolium* LA1589 recombinant inbred backcross lines (IBLs) (Doganlar et al. 2002), in parallel with an $F_2$ mapping population derived from LA1589 and the susceptible cultivar Heinz BG-1706 to identify genomic regions and candidate genes associated with resistance to *Pst*19. The identification of these regions will facilitate the future breeding of varieties resistant to bacterial speck caused by *Pst* race 1 strains.

## Materials and Methods

*Plant materials and growth conditions*

Tomato seeds were sterilized in 50% bleach for 30 min. After sterilization, the seeds were rinsed five times with sterile nanopure $H_2O$ and then plated or sown out on soil. Seeds for plate experiments were germinated on 100 x 25 mm plates containing sterile 0.5 X Murashige and Skoog (MS) basal salts and 0.8% agar. Seeds used for soil experiments were planted in Sunshine Mix#1/LC1 (Sun Gro Horticulture Canada Ltd.) supplemented with 15:9:12 fertilizer. Seeds on plates or soil were stratified at 4°C for 3 days to synchronize germination. Plants were grown in a growth chamber under a constant temperature of 22ºC and 16 h of light (200-220 $\mu E\ m^{-2}\ s^{-1}$) and 8 h of darkness. The following tomato accessions were obtained from the Tomato Genetics Resource Center (tgrc.ucdavis.edu): *S. lycopersicum* MoneyMaker-*PtoS*, *S. lycopersicum* MoneyMaker-*PtoR*, *S. lycopersicum* LA3342 (RioGrande-*PtoR*), *S. lycopersicum* LA3343 (RioGrande-*PtoS*), *S. lycopersicum* LA4345 (Heinz BG-1706), *S. pimpinellifolium* LA1589 (also called PI407545), *S. lycopersicum* LA4024 (E-6203) and recombinant Inbred Backcrossed Lines (IBL) *S. pimpinellifolium* LA4139 – LA4229 (Doganlar et al. 2002). The IBLs originally consisted of 196 lines of which 100 were identified based on uniform genome coverage and map resolution (Doganlar et al. 2002). 87 lines were available from the Tomato Genetics Resource Center.

*P. syringae strains, culture conditions and infection assays*

*P. syringae* pv. *tomato* strains were grown in King's broth (KB). Antibiotics were used at the following concentrations: 50 µg/mL rifampicin dissolved in dimethylformamide, 50 µg/mL cycloheximide dissolved in ethanol.

For seedling hypersensitive response (HR) assays, *Pst*19 was resuspended to an optical density at 600 nm ($OD_{600}$) of 0.1 (approximately 5 x $10^7$ CFU/ml) and pressure infiltrated into both cotyledons. The HR was scored 16-20 h post-infiltration. For the seedling flood assay, *Pst*19 and *Pst*DC3000 were resuspended to an $OD_{600}$ of 0.1 and then serially diluted in 10 mM $MgCl_2$ to a final $OD_{600}$ of 0.0075 and 0.005, respectively with 0.015% Silwet L-77 (Hassan et al. 2020). Ten-day-old seedlings were flooded for 3 min with 6 mL of inoculum or 10 mM $MgCl_2$.

Seedlings were phenotyped for disease or resistance 10-14 days after flooding. For bacterial growth assays in seedlings, seedlings were flooded as described above and four days later, one cotyledon was removed, surface sterilized in 70% ethanol for 10 s and rinsed in nanopore $H_2O$ for 10 s. Each cotyledon was blotted, individually weighed and homogenized in 10 mM $MgCl_2$. Homogenized lysate was plated on KB with 50 µg/mL rifampicin and 50 µg/mL cycloheximide for colony counting. Cycloheximide prevents fungal contamination. Colony counts for seedlings were normalized to 0.1 g of tissue for cotyledons (Hassan et al. 2017).

For infection assays in adult plants, *P. syringae* was resuspended in 10 mM $MgCl_2$ to an $OD_{600}$ of 0.2 with 200 µL/L of Silwet L-77. Adult plants at the 4-6 leaf stage were inverted, then submerged and swirled in the inoculum for 30 s. Infected plants were incubated in a growth chamber under a humidity dome for 1-2 days at which time the dome was removed. Adult bacterial growth assays were performed 5 days after infection on abaxial leaflets of the 4$^{th}$ leaf. Tissue was sterilized as described above for seedlings. A total of 1 $cm^2$ tissue (four disks) was harvested from infected leaves, ground in 10 mM $MgCl_2$, and plated as described for seedlings. Plants were re-incubated and phenotyped 7 days after infection.

*Screening and analysis of IBL population*

IBLs were screened in the seedling flood assay as described above, and the number of surviving or deceased individuals was counted. Selected IBLs were tested for bacterial growth and symptoms in adult plants as described above.

Single-marker analysis was performed on the IBLs in the seedling flood assay using the Wilcoxon Mann-Whitney test, a nonparametric counterpart of the t-test because this test does not assume the probability distribution of the quantitative trait. For each marker, survival rates of seedlings homozygous for either parental allele were compared to determine significant differences. Marker loci were determined to be highly significantly associated with resistance at $p<0.01$, significantly associated with resistance at $p<0.05$ and suggestive of resistance at $p<0.1$. QTL mapping intervals were defined as regions including marker loci significantly associated with resistance, whose boundaries were defined at the first instance of an adjacent nonsignificant marker locus.

*Generation and screening the F2 Heinz-BG1706 x S. pimpinellifolium LA1589 mapping population*

Heinz 1706-BG and a *Pst*19-resistant S. *pimpinellifolium* LA1589 individual were crossed. Seedlings from the F2 segregating population were grown on plates and flooded using the seedling flood assay protocol (Hassan et al. 2020). Each seedling was labeled with a unique number. Two days after flooding, the tip of one cotyledon from each seedling was snipped, frozen in liquid $N_2$ and stored at -80ºC. Plates were resealed and seedlings re-incubated in the growth chamber at 22ºC. Highly susceptible seedlings were phenotyped 7-9 days after infection and strongly resistant seedlings were phenotyped 10 days after infection.

*Nucleic acid isolation, library construction, and sequencing*

Genomic DNA was extracted from pools of strongly resistant LA1589 or highly susceptible LA1589 F2 individuals, respectively using the Puregene Core Kit A (Qiagen Inc.). Libraries were constructed at the Functional Genomics Library (FGL), a QB3-Berkeley Core Research Facility at UC Berkeley. A S220 Focused-Ultrasonicator (Covaris) was used to fragment genomic DNA to 1-6000 bp and library preparation was performed using the KAPA

Hyper Prep kit for DNA (KK8504). Truncated universal stub adaptors were used during PCR amplification to complete the adapters and to enrich the libraries for adapter-ligated fragments. Samples were checked for quality on an AATI Fragment Analyzer. Samples were then transferred to the Vincent J. Coates Genomics Sequencing Laboratory (GSL), another QB3-Berkley Core Research Facility at UC Berkeley, where Illumina sequencing library molarity was measured with quantitative PCR with the Kapa Biosystems Illumina Quant qPCR Kits on a BioRad CFX Connect thermal cycler. Libraries were then pooled evenly by molarity and sequenced on an Illumina NovaSeq6000 150PE S4 flowcell, generating 10 Gb minimum of data per sample. Raw sequencing data was converted into fastq format, sample specific files using Illumina bcl2fastq2 software.

*Quality control, alignment, and variant calling*

Paired-end Illumina read quality was checked with FastQC v0.11.7 (Andrews 2010) and read fastq files were trimmed using Cutadapt v2.4 (Martin 2011) to remove TruSeq adapter sequences 5'ACACTCTTTCCCTACACGACGCTCTTCCGATCT3' and 5'GATCGGAAGAGCACACGTCT3'. BWA-MEM v0.7.17 (Li and Durbin 2009; Li 2013) was used to align paired-end reads from strongly resistant LA1589 individuals or highly susceptible LA1589 individuals to: 1) the SL4.0 Heinz BG-1706 reference build (Su et al. 2021), 2) the *Solanum pimpinellifolium* LA2093 genome (Wang et al. 2020) and, 3) the *Solanum pimpinellifolium* LA1589 PacBio genome (Alonge et al. 2020). Variants in candidate genes were identified by mapping to Heinz BG-1706, LA1589 or LA2093. Heinz BG-1706 variants were identified using Wgsim (Li 2021) to simulate $10^8$ Illumina reads from the Heinz BG-1706 genome (SL4.0 build). Simulated reads from Heinz BG-1706 were aligned to the LA1589 Pacbio reference genome (Alonge et al. 2020) or the LA2093 reference genome (Wang et al. 2020) using BWA-MEM v0.7.17. BAM files from the alignments were sorted using SAMtools v1.8 (Li et al. 2009). BCFtools v.16 was used to perform variant calling. Low-quality variants with read depths of less than 10 were filtered out.

SnpSift was used to further filter candidate genes for homozygous variants (Cingolani et al. 2012). Variants with a maximum fraction of reads supporting an indel of less than 0.1 were filtered out (http://www.htslib.org/doc/). SnpEff databases were built using SL4.1, LA2093 PacBio and LA1589 Pacbio annotations (Alonge et al. 2020; Wang et al. 2020; Su et al. 2021). SnpEff was used to predict the effects of variants on the translation of annotated genes (Cingolani et al. 2012). Genes were considered candidates for resistance or susceptibility if a) they were within the mapping intervals from the IBL screen and b) had variants for which the SNPEff putative variant impact was classified as high or moderate impact. SNPEff predicts high impact variants to have a disruptive effect on the protein and moderate impact variants to have changes that are not disruptive but may alter the effectiveness of the protein.

LA1589 orthologs of Heinz BG-1706 or LA2093 candidates were identified using Reciprocal best Basic Local Alignment Search Tool (BLAST version 2.2.31+). Genome-to-genome protein sequence comparisons were made between the query genome (LA1589) and the database genome (Heinz BG-1706 SL4.1), as well as the reciprocal sequence comparisons. The LA1589 and LA2093 proteins were compared in the same manner. A maximum E-value threshold of $1x10^{-6}$ was used with Smith-Waterman alignment (Ward and Moreno-Hagelsieb 2014). Top best hits for query proteins were sorted and compared to top best hits in the reciprocal direction. If no orthologs were identified, then closest homologs were identified using BLAST version 2.2.31+ for high bit-score and low e-values (Pearson 2013).

## Results

*S. pimpinellifolium LA1589 exhibits resistance to Pst19 in both tomato seedlings and adult plants*

We previously screened 96 wild tomato accessions for *Pst*19 resistance using a high-throughput seedling resistance assay (Hassan et al. 2017). In subsequent screens, we identified an additional wild accession, *S. pimpinellifolium* LA1589, with *Pst*19 resistance. LA1589 seedlings displayed resistance to infection (28/34, 82%) whereas susceptible RioGrande-*PtoR* (RG-*PtoR*) seedlings died (0/9, 100%) (Figure 1A). To quantitatively confirm our phenotypic observations, we carried out bacterial growth assays on LA1589 (n=23) and Moneymaker-*PtoS* (MM-*PtoS*) (n=10) seedlings flooded with *Pst*19. *Pst*19 grew to log 7 in LA1589 seedlings (n=23) compared to log 8 growth in MM-*PtoS* seedlings (n=10) (Figure 1B). Therefore, LA1589 supported 1 log less bacterial growth compared to the susceptible MM-*PtoS* cultivar.

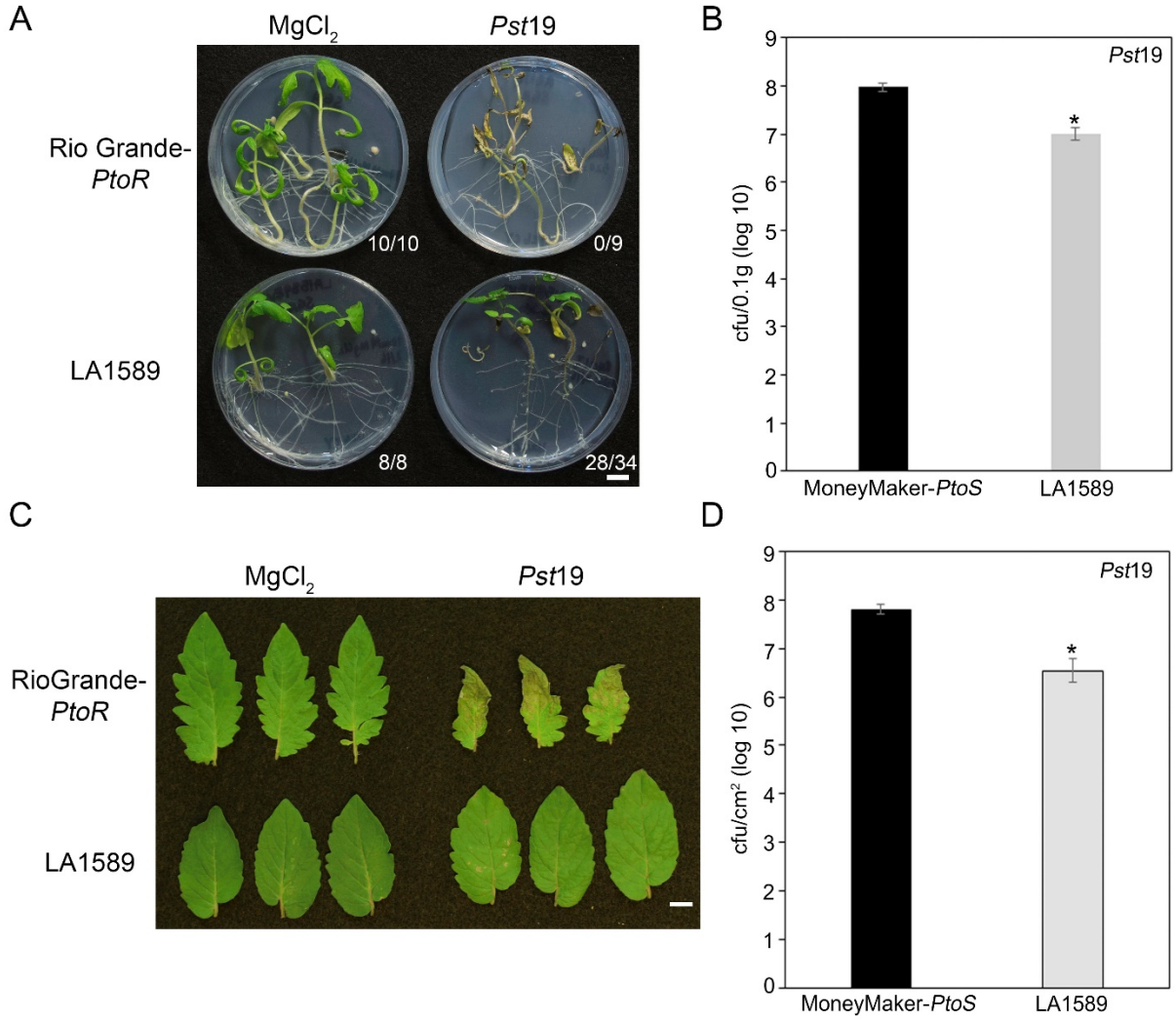



**Figure 1.** Both LA1589 seedlings and adults exhibit qualitative and quantitative resistance to *Pst*19. (A) Phenotypic resistance or disease symptoms in Rio Grande-*PtoR* or LA1589 tomato seedlings 10-14 days after being flooded with *Pst*19 at an $OD_{600}$ of 0.0075. The number of surviving plants is indicated over the total number of tested plants. (B) Bacterial counts were determined 4 days post-infection (dpi) on MoneyMaker-*PtoS* (n=10) and LA1589 (n=23) seedlings and normalized to 0.1 g of tissue. The asterisk above the bar indicates statistical difference as determined

by a one-factor ANOVA using a GLM procedure ($p<0.05$). The error bars indicate the standard error. The experiment was repeated 3 times with similar results. (C) Rio Grande-*PtoR* or LA1589 were grown to the 4-6 leaf stage, and dip-inoculated with a bacterial suspension at an $OD_{600}$ of 0.2. Leaves were photographed 7 days after infection. The scale bar indicates 1 cm. (D) Bacterial counts were determined 5 days post infection on LA1589 (n=10) and MoneyMaker-*PtoS* (n=5) plants. The asterisk above the bar indicates statistical difference as determined by a one-factor ANOVA using a GLM procedure ($p<0.05$). The error bars indicate the standard error. The experiment was repeated three times with similar results.

To determine whether resistance to *Pst*19 is maintained in adult plants, we carried out dip inoculations with RG-*PtoR* plants grown to the 4- to 6 leaf stage and assessed the plants for disease symptoms 7 days past infection (dpi). We previously confirmed that MM-*PtoS*, MM-*PtoR*, RG-*PtoS* and RG-*PtoR* are suitable as controls since all are susceptible to *Pst*19 and show similar severe disease symptoms (Hassan et al. 2017). LA1589 exhibited very few specks on leaves and very mild disease symptoms, compared to the control, RG-*PtoR* (Figure 1C). RG-*PtoR* displayed numerous lesions and leaf collapse. To quantitatively confirm these phenotypic observations, we measured bacterial growth in LA1589 plants. LA1589 (n=10) supported log 6.5 growth of *Pst*19, approximately 1.3 log less growth than that observed in MM-*PtoS* (n=5, log 7.8) at 5 dpi (Figure 1D). Taken together, these qualitative and quantitative results support resistance in LA1589 seedlings and adult plants to *Pst*19.

*S. pimpinellifolium LA1589 displays Pto-dependent and Pto-independent resistance to Pseudomonas syringae pv. tomato*

*S. pimpinellifolium* was the original source of the *Pto/Prf* cluster (*PtoR*), which confers resistance to race 0 strains (Pitblado and Macneill 1983; Pedley and Martin 2003). To confirm that LA1589 can recognize a race 0 strain, we flood inoculated seedlings with *Pseudomonas syringae pv. tomato* DC3000 (*Pst*DC3000). As expected, all LA1589 seedlings (n=17) and the positive control line, Moneymaker-*PtoR* (MM-*PtoR*) (n=14), were resistant to *Pst*DC3000 (Figure 2A). We also tested Heinz BG-1706 which does not carry the *Pto/Prf* gene cluster, and found that all Heinz BG-1706 seedlings (n=12) were susceptible to *Pst*DC3000 (Figure 2A). Therefore, the Pto/Prf cluster is functional in LA1589, allowing the recognition of effectors from *Pst*DC3000.

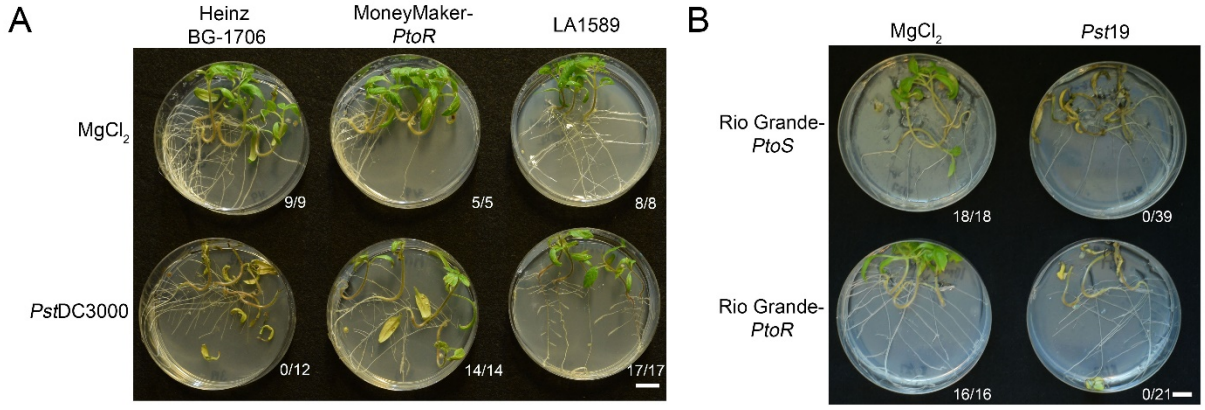



**Figure 2.** LA1589 seedlings display *Pto*-dependent resistance to *Pst*DC3000. (A) Phenotypic resistance or disease symptoms in Heinz 1706-BG, MoneyMaker-*PtoR* and LA1589 seedlings 7-10 days post-infection with *Pst*DC3000 at an $OD_{600}$ of 0.005. Heinz-BG1706 seedlings lack the *Pto*/*Prf* gene cluster and are susceptible to *Pst*DC3000. MoneyMaker-*PtoR* and LA1589 seedlings carry the *Pto*/*Prf* gene cluster and are resistant. The number of surviving seedlings out of the total number tested is shown under each plate. The scale bar is 1 cm. (B) Rio Grande seedlings with (*PtoR*) and without (*PtoS*) the *Pto/Prf* gene cluster are susceptible to *Pst*19 flooded at an $OD_{600}$ of 0.0075.

Disease symptoms in Rio Grande-*PtoS* and Rio Grande-*PtoR* seedlings are shown 7-10 days post-infection. The number of surviving seedlings out of the total number tested is shown under each plate. The scale bar is 1 cm.

While *Pst*DC3000 and *Pst*19 both infect tomato, *Pst*19 is closely related to *Pst*T1, which is a distinct strain with a unique effector complement compared to *Pst*DC3000 (Almeida et al. 2009; Kunkeaw et al. 2010). *Pst*T1 and *Pst*19 do not carry *avrPto* and do not appear to accumulate AvrPtoB (Lin et al. 2006; Kunkeaw et al. 2010). To determine whether *Pst*19 might express a low level of *avrPtoB* which could be recognized through Pto/Prf in LA1589, we carried out seedling flood assays on RG-*PtoR* (n=21) which contains the *Pto/Prf* locus and RG-*PtoS* (n=39) which lacks the *Pto/Prf* locus. All seedlings died in both genotypes, regardless of whether the lines had or lacked the *Pto/Prf* cluster. This indicates that resistance to *Pst*19 in LA1589 is independent of *Pto/Prf* (Figure 2B). This result is consistent with previous work showing that AvrPtoB in *Pst*19 is not recognized by *S. habrochaites* LA2109 (Bao et al. 2015).

*Pst19 does not elicit a hypersensitive response characteristic of ETI*

To determine whether recognition of a T3SE in *Pst*19 may cause ETI, we investigated whether *Pst*19 can trigger an HR in LA1589. We infiltrated 10-day old seedlings of LA1589 or MM-*PtoS* (which lacks the *Pto/Prf* locus) with *Pst*19, as previously described (Hassan et al. 2017). *Pst*19 did not trigger an HR in LA1589 (0/16) or MM-*PtoS* seedlings (0/16) (Figure 3). This suggests that *Pst*19 triggers non-canonical resistance in LA1589, which is consistent with the observed levels of bacterial growth (Figure 1B).

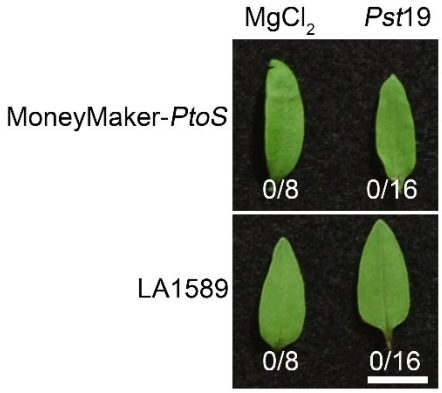



**Figure 3.** Resistance to *Pst*19 in LA1589 is not associated with a hypersensitive response. 10 day-old Moneymaker-*PtoS* (MM-*PtoS*) or LA1589 seedlings were pressure infiltrated with ~ 5.0 x $10^7$ cfu/mL *P. syringae* pv. *tomato* 19 (*Pst*19). Plants were photographed 18-22 hours after infiltration. The number of leaves exhibiting an HR is shown under the leaves. The scale bar is 1 cm.

*QTL mapping using S. pimpinellifolium LA1589 inbred backcross lines identifies five genomic regions associated with Pst19 resistance*

To identify genomic regions associated with resistance in LA1589, we took advantage of a previously generated inbred backcross line (IBL) population from a cross between *S. lycopersicum* E-6203 and *S. pimpinellifolium* LA1589 (Doganlar et al. 2002). The IBLs were generated after two backcrosses and 6 generations of inbreeding via single seed descent ($BC_2F_6$) and are highly homozygous (Doganlar et al. 2002). Of the original 100 IBLs with uniform genome coverage and map resolution (Doganlar et al. 2002), we were able to obtain 87 lines from the Tomato Genetics Resource Center.

We screened approximately 20 individuals for each of 87 homozygous $BCF_6$ lines in the seedling assay. For 26 IBLs, at least one seedling survived 10-14 dpi (Table S1). We identified 4 IBLs with 30-49% seedling survival (Group 1), 9 IBLs with 10-19% survival (Group 2) and 13 IBLs with 1-9% seedling survival (Group 3) (Table S1). *Pst*19-resistant seedlings displayed a healthy shoot apical meristem and new green vegetative growth. For 61 IBLs, all seedlings were susceptible to *Pst*19 infection (Table S1). Susceptible seedlings were dead and displayed brown apical meristems and a lack of new growth. The recurrent parent E-6203 consistently died when flooded with *Pst*19 as a susceptible control in the screen (Figure S1A).

To detect genetic associations with *Pst*19 resistance, we undertook single marker analysis based on the frequency of seedling survival and the presence of restriction fragment length polymorphism (RFLP) markers homozygous for LA1589. Five QTL were identified and were named quantitative resistance to *Pst*19 in *S. pimpinellifolium* (qRpp1-5, qRpp1-6, qRpp1-8.1, qRpp1-8.2, qRpp1-12). Markers with a p-value<0.01 were considered highly significant, markers with a p-value<0.05 were considered significant, markers with a p-value<0.01 were considered suggestive. qRpp1-5 contains one significant marker CT101 ($p<0.05$), and one suggestive marker TG441 ($p<0.1$) on chromosome 5 (Table 1, Figure 4). The next marker on chromosome 5, CT167, was not linked with resistance and was used to delineate the resistance interval of ~1.8 MB (Figure 4). qRpp1-6 contains one significant marker CT216 ($p<0.05$) (Table 1, Figure 4) on chromosome 6. The next marker on chromosome 6, TG178, was not linked with resistance and the resistance interval is ~23.4 MB. qRpp1-8.1 has one suggestive marker TG176 ($p<0.1$) on the top of chromosome 8. The next marker with sequence was TG349, which is quite far from TG176. The interval from the tip of the chromosome to TG349 is ~55.4MB. qRpp1-8.2 has one highly significant marker TG201 ($p<0.01$) and one significant marker CT265 ($p<0.05$) at the bottom of chromosome 8 (Table 1, Figure 4). TG201 had the greatest significance of all the markers and is associated with the highest phenotypic effect. The qRpp1-8.2 interval on chromosome 8 is ~59.1 MB, and is delineated by two unlinked markers, CT302 and CT68. The CT302 marker is quite distant from the markers of interest but had to be used because TG330 and TG505 lacked genotypic information. Finally, qRpp1-12 contains one suggestive marker TG565 ($p<0.1$) on chromosome 12 (Table 1, Figure 4). The adjacent markers, CT99 and TG111, are not linked with resistance, which delineates a region of ~56.6 MB.

**Table 1.** Identification of molecular markers linked to *Pst*19 resistance in LA1589 using single marker analysis

| QTL name | Marker | Chr[1] # | P-value[2] |
|---|---|---|---|
| qRpp1-5 | CT101** | 5 | 0.04 |
| | TG441* | 5 | 0.06 |
| qRpp1-6 | CT216** | 6 | 0.03 |
| qRpp1-8.1 | TG176* | 8 | 0.07 |
| qRpp1-8.2 | TG201*** | 8 | 0.006 |
| | CT265** | 8 | 0.02 |
| qRpp1-12 | TG565* | 12 | 0.08 |

[1]Chr indicates chromosome

[2]Significance of markers determined using Wilcoxon Mann-Whitney test and the following p-values.

***Highly significant markers ($p<0.01$)

**Significant markers (p<0.05)
*Suggestive markers (p<0.1)

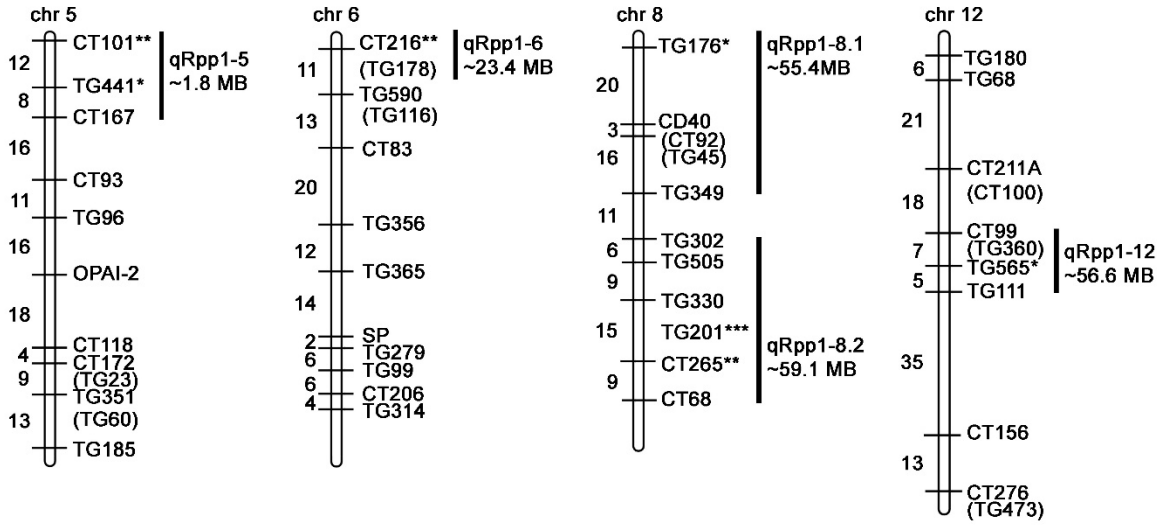



**Figure 4.** Schematic of markers on chromosomes 5, 6, 8, and 12 in the IBL population indicating QTL associated with resistance (qRpp). Map is modified from (Doganlar et al. 2002) and https://solgenomics.net/cview/map.pl?map_version_id=26. Regions encompassing highly significant markers (*** p<0.01), significant markers (** p<0.05) and/or suggestive markers (* p<0.1) are indicated. Adjacent markers that are not linked with resistance delineate the maximum region of interest (Table S2). Some marker sequences were not available (Table S3). The distance between markers was determined based on blasting the Heinz BG-1706 genome with the marker sequences (Table S3).

To refine the genomic intervals associated with resistance, we mapped relevant markers onto the genetic map by blasting the Heinz BG-1706 genome (build 4.0) with marker sequences. We were able to identify genomic sequences for all significant, suggestive or adjacent markers except CD40, CT111, TG330, TG505, TG201 and CT265 (all on chromosome 8) which lacked sequence information (Table S2, S3). The intervals identified in the IBL analysis contained 244 genes on chromosome 5, 586 genes on chromosome 6, 1335 or 1027 genes in the two regions on chromosome 8, and 1148 genes on chromosome 12 (Table 2, Figure 4).

**Table 2.** Number of genes and loci in LA1589 delineated by IBL markers and associated with resistance to *Pst*19

| **Chr**[1] | **Delineated markers** | **# of genes** | **Loci in LA1589** |
|---|---|---|---|
| 5 | 0-CT167 | 244 | Spim05g005010-Spim05g007440 |
| 6 | 0-TG178 | 586 | Spim06g005010-Spim06g010850 |
| 8 | 0-TG349 | 1335 | Spim08g005010-Spim08g018350 |
| 8 | TG302-CT68 | 1027 | Spim08g020700-Spim08g030960 |
| 12 | TG360-TG111 | 1148 | Spim12g011350-Spim12g022830 |

[1]Chr indicates chromosome

Based on the single marker analysis, we selected IBLs with various combinations of the four highly significant, significant or suggestive markers (Table 1, S4). Previous analysis of the IBLs had determined whether the markers were homozygous for LA1589, homozygous for E-

6203, or heterozygous with both LA1589 and E-6203 (Table S4, Figure 5) (Doganlar et al. 2002). No IBLs were homozygous for LA1589 at all four highly significant or significant markers associated with the QTL. We prioritized testing seven IBLs because they showed higher levels of resistance (Table S1), 4 from Group 1 with 30-49% resistance (LA4156, LA4168, LA4208, LA4216) and 3 from Group 2 with 10-19% resistance (LA4144, LA4148, LA4173). We dip-inoculated adult plants from LA1589, the seven IBLs with higher levels of resistance, a susceptible IBL (LA4152), and the susceptible E-6203 cultivar with *Pst*19, and assessed the development of disease symptoms. Typical symptoms of bacterial speck include brown or dark brown necrotic lesions surrounded by chlorotic halos. All seven of the resistant IBLs (LA4144, LA4148, LA4156, LA4168, LA4173, LA4208, LA4216) displayed reduced lesions, compared to the susceptible lines E-6203 and LA4152. However, the IBLs displayed more lesions than LA1589, which had the fewest lesions (Figure 5).

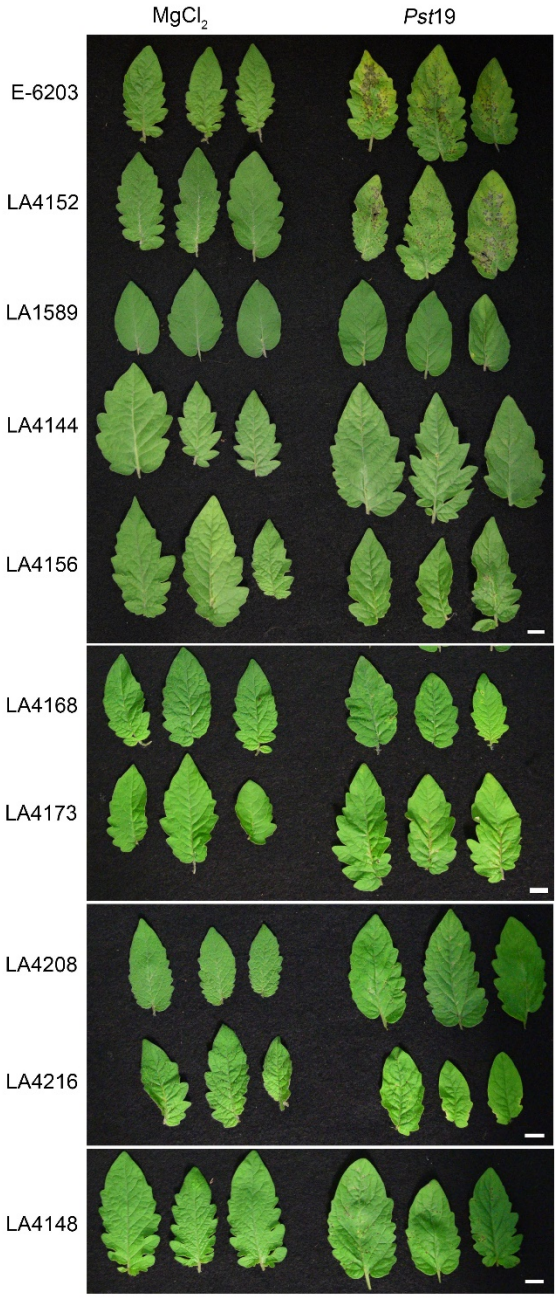



**Figure 5**. Inbred backcross lines carrying markers associated with *Pst*19 resistance show less disease to *P. syringae* pv. *tomato* 19 (*Pst*19) than a susceptible cultivar. Susceptible cultivar, E-6203, susceptible IBL LA4152, resistant LA1589 and seven IBLs carrying markers associated with resistance (LA4144, LA4156, LA4168, LA4173, LA4208, LA4216 and LA4148) were dip-inoculated with a bacterial suspension at an $OD_{600}$ of 0.2 at the 4-6-week-old adult stage. Leaves were photographed 6-7 days after infection. The scale bar indicates 1 cm.

To quantitatively confirm these phenotypic results, we measured bacterial growth in all seven IBLs and compared them to the susceptible E-6203 cultivar and susceptible IBL LA4152, as well as the resistant accession LA1589. Five IBLs, LA4144 (log 6.6), LA4156 (log 7.1), LA4168 (log 6.7), LA4173 (log 7.3) and LA4216 (log 7.3) displayed between an 8-to-10-fold reduction in bacterial growth, compared to the parental line, E-6203 and a 7-to-13-fold reduction,

compared to the susceptible IBL LA4152 (Figure 6). One IBL, LA4208 (log 7.8) displayed a more modest but still significant 3-fold reduction in growth, compared to E-6203 and LA4152 (Figure 6). LA4148 supported similar levels of bacterial growth as LA4208 but these differences were not statistically significant compared to E-6203 and LA4152 (Figure 6). Bacterial growth in all IBLs was, however, significantly higher than in LA1589, ranging from 0.7-1.9 higher across these lines. These results are consistent with the intermediate symptoms observed in these lines as compared to LA1589 (Figure 5). The combination of markers within the lines contribute to a partial reduction in bacterial growth, but do not recapitulate the complete reduction in bacterial growth observed in LA1589.

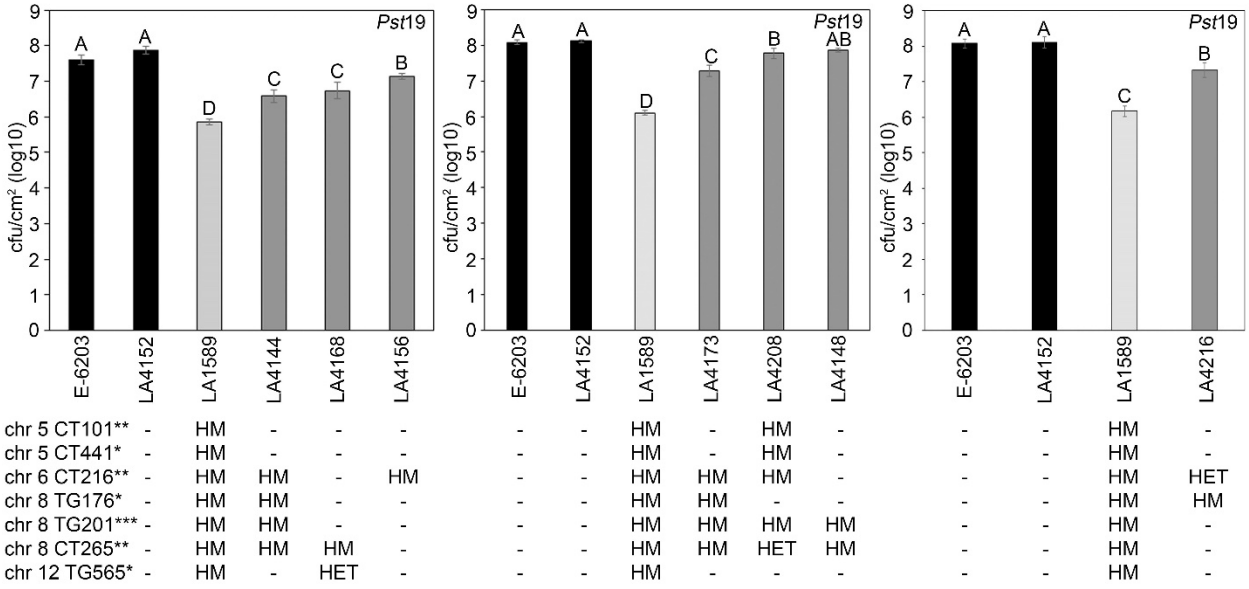



**Figure 6**. Inbred backcross lines carrying markers associated with *Pst*19 resistance support less growth of *P. syringae* pv. *tomato* 19 (*Pst*19). Susceptible cultivar E-6203 (black bar), susceptible IBL LA4152 (black bar), resistant LA1589 (light grey bar) and seven IBLs carrying markers associated with resistance LA4144, LA4148, LA4156, LA4168, LA4173, LA4208 and LA4216 (dark grey bars) were dip-inoculated with a bacterial suspension at an $OD_{600}$ of 0.2 at the 4-6-week-old adult stage. Bacterial counts and log growth were determined at 5 days past infection (dpi) in E-6203, LA4152, LA1589, LA4144, LA4168 and LA4156 (left panel), E-6203, LA4152, LA1589, LA4173, LA4208 and LA4148 (middle panel), and E-6203, LA4152, LA1589 and LA4216 (right panel). Within each experiment, significant differences among genotypes are shown with different letters. One-factor ANOVA using a GLM procedure ($p<0.05$) followed by the Tukey's *post hoc* test, a multiple comparison of means test, was used to determine statistical differences. The error bars indicate standard error. Experiments for each IBL were repeated three times with similar results. The genotypes of the IBLs are shown below the growth assay with the chromosome location and marker genotype for each highly significant (*** $p<0.01$), significant (** $0.01<p<0.05$), or suggestive (* $0.05<p<0.1$) marker (see also Table 1, S4). HM indicates homozygous LA1589. HET indicates heterozygous LA1589/E-6203. - indicates homozygous E-6203. The genotype of LA1589 is inferred. qRpp1-5 includes CT101 and TG441. qRpp1-6 includes CT216. qRpp1-8.1 includes TG176 and qRpp1-8.2 includes TG201 and CT265. qRpp1-12 includes TG565.

*F2 mapping population approach identifies candidate genes for resistance*

To refine the genetic intervals from the IBLs and identify SNPs associated with resistance, we generated an F2 mapping population from a cross between LA1589 and the *Pst*19-susceptible cultivar Heinz BG-1706 (Figure S1A). We infected and phenotyped 1181 plants for resistance or susceptibility to *Pst*19, using the previously described criteria. We found that 18% (n=215) were resistant and 82% (n=966) were susceptible (Table 3). The segregation ratios are

not Mendelian, consistent with the IBL data showing that multiple loci contribute to resistance (Table 1). To further refine the resistance or susceptibility phenotypes, we identified five phenotypic groups as follows: strongly resistant seedlings (6.7%), moderately resistant seedlings (3.7%), weakly resistant seedlings (7.8%), susceptible seedlings (62%) and highly susceptible seedlings (20%) (Table 1). Strongly resistant seedlings displayed the greatest degree of branching and new growth compared to other seedlings at 14 dpi. Moderately resistant seedlings had some branching and new growth, and weakly resistant seedlings had very little branching and one or two newly emerging leaves. Highly susceptible seedlings had brown apical meristems and no new green growth at 7-9 dpi, whereas susceptible seedlings took longer to exhibit these phenotypes. To maximize the identification of potential genetic differences contributing to the resistance phenotype, we selected individuals at the phenotypic extremes: 79/1181 (6.7%) seedlings with strong resistance, and 233/1181 (20%) with high susceptibility. We extracted DNA from a pool of 79 seedlings with strong resistance, and a pool of 233 highly susceptible individuals, and carried out whole-genome sequencing. We obtained 348-397X coverage of the genome, where 73-74% of reads for the strongly resistant pool and 65-67% of reads for the highly susceptible pool had a BWA MEM quality alignment score of 20 or higher (Table S5).

**Table 3.** *Pst*19 resistance in F2 segregating population of Heinz BG-1706 x LA1589 seedlings

| **Phenotype** | **Number of plants (percentage)** | **Subcategory** | **Number of plants (percentage)** |
|---|---|---|---|
| Susceptible | 966 (82%) | Highly susceptible | 233 (20%) |
| | | Susceptible | 733 (62%) |
| Resistant | 215 (18%) | Strongly resistant | 79 (6.7%) |
| | | Moderately resistant | 44 (3.7%) |
| | | Weakly resistant | 92 (7.8%) |
| Total | 1181 | | 1181 |
| Chi-squared goodness of fit test followed by Yates correction confirms that the hypothesis of a Mendelian segregation ratio of 3 susceptible:1 resistant with an error rate of 5% could not be accepted ($p<0.001$). | | | |

We mapped reads from the strongly resistant and highly susceptible pools to the Heinz BG-1706, *S. pimpinellifolium* LA2093 and *S. pimpinellifolium* LA1589 reference genomes (Sato et al. 2012; Alonge et al. 2020; Wang et al. 2020) and identified variants (Figure 7). For reads mapped to Heinz BG-1706, we assumed that SNPs from LA1589 were important for resistance, and therefore selected variants specifically present in the strongly resistant pool and absent in Heinz (Figure 7, left, outlined in bold black). To eliminate variants commonly found in *S. pimpinellifolium* that do not contribute to resistance, we mapped reads to the susceptible *S. pimpinellifolium* LA2093 genome and looked for variants unique to the strongly resistant pool (Figure 7, S1B). Since variants in the strongly resistant pool mapped to LA2093 may come from the cross with Heinz BG-1706, we removed variants specific to Heinz BG-1706. This enabled us to identify unique variants specific to LA1589 which may contribute to resistance (Figure 7, middle, outlined in bold black). Lastly, for reads mapped to *S. pimpinellifolium* LA1589, unique SNPs in the highly susceptible pool could come from either parent and might disrupt resistance. These SNPs are of interest if they differ between the strongly resistant and highly susceptible

pools (Figure 7, right, blue bubble outlined in bold black). We also assumed that a small number of SNPs could be unique to the strongly resistant pool since there could be some diversity between the sequenced LA1589 and the TGRC stock of LA1589 (Figure 7, right, green bubble outlined in bold black). Variants relative to LA1589 that are present in Heinz BG-1706 and the strongly resistant pool are not useful and were eliminated. Based on all three approaches, we identified 11 candidate genes, with unique variants in the resistant pool compared to the susceptible pool.

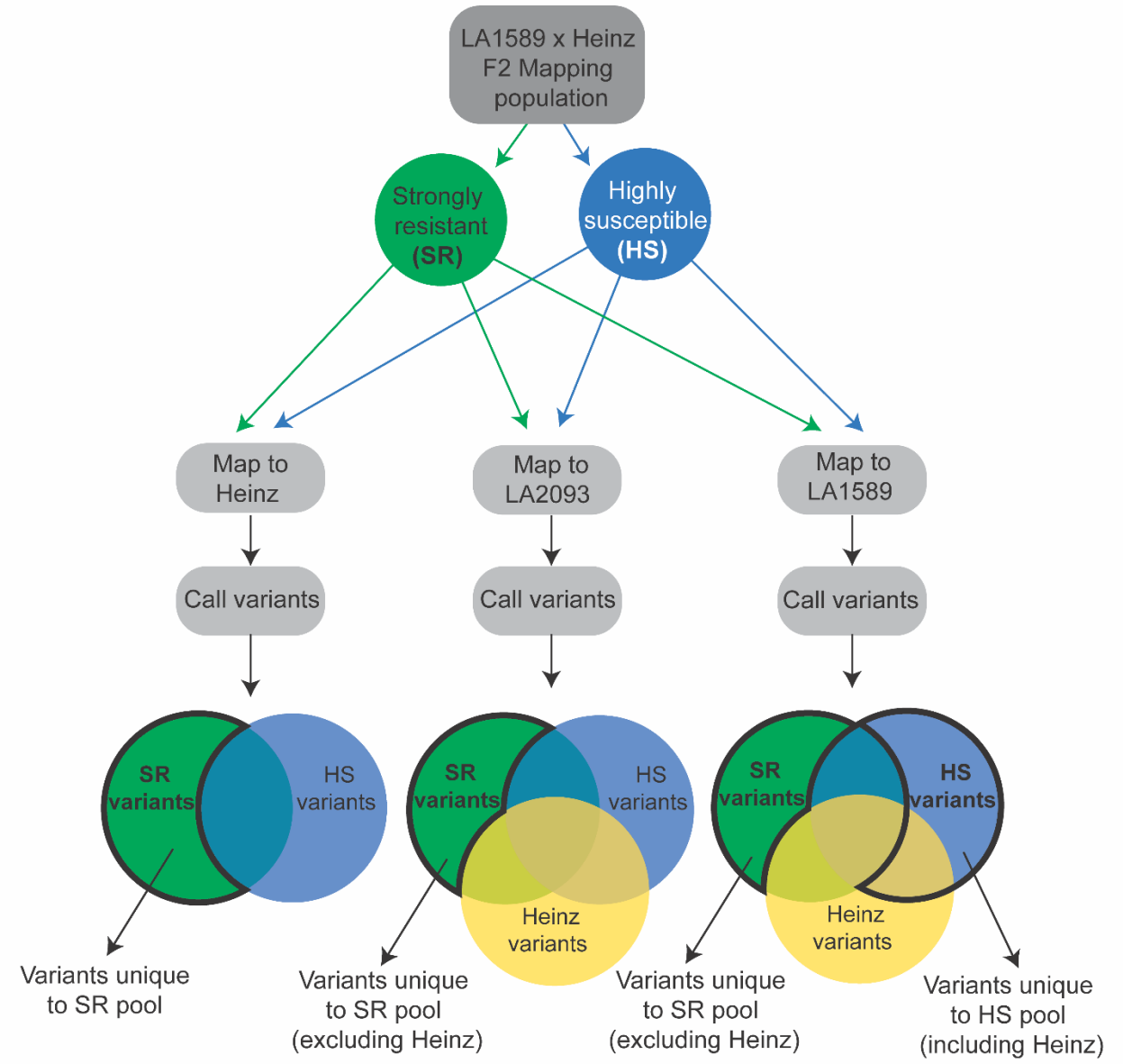



**Figure 7.** Schematic of variant calling pipeline to susceptible Heinz BG-1706, susceptible *S. pimpinellifolium* LA2093 and resistant *S. pimpinellifolium* LA1589. SR indicates strongly resistant (green) and HS indicates highly susceptible (blue). Heinz variants are shown in yellow. Variants of interest in the Venn diagrams are shown with a bold black outline.

qRpp1-6 contains one candidate gene: a SUMO-domain containing protein. qRpp1-8.1 contains three candidate genes: conserved oligomeric Golgi (COG) complex subunit 2, a

hypothetical protein, and an AT-rich interactive domain-containing protein 5-like (Table 4). qRpp1-8.2 contains four candidate genes: a protein phosphatase 2C (PP2C), a pectin acetylesterase, a copper chaperone for superoxide dismutase, and ribosomal protein L16. qRpp1-12 contains three candidate genes: a NADH-ubiquinone oxidoreductase chain protein, a zinc finger CCCH-domain containing protein and a hypothetical protein. No candidate genes were identified for qRpp1-5.

**Table 4.** Candidate genes associated with resistance from IBL and GBS analysis

| QTL | Gene number[1] | Annotation | Identified by mapping to: |
|---|---|---|---|
| qRPP1-6 | SPIMP06g0183860 | SUMO-domain containing | LA2093 |
| qRPP1-8.1 | Solyc08g006660 | COG2 | Heinz |
| qRPP1-8.1 | Spimp08g006910 | Hypothetical | LA1589 |
| qRPP1-8.1 | SPIMP08g0244070 | AT-rich interactive domain-containing protein 5-like | LA2093 |
| qRPP1-8.2 | Spimp08g024080 | Protein phosphatase 2C | LA1589 |
| qRPP1-8.2 | SPIMP08g0250550 | Pectin acetylesterase | LA2093 |
| qRPP1-8.2 | Solyc08g079830 | Copper chaperone for superoxide dismutase | Heinz |
| qRPP1-8.2 | Solyc05g045840 | Ribosomal protein L16 | Heinz |
| qRPP1-12 | Spimp12g015200 | NADH-ubiquinone oxidoreductase chain | LA1589 |
| qRPP1-12 | Spimp12g015620 | Zinc finger CCCH domain-containing protein | LA1589 |
| qRPP1-12 | Spimp12g018680 | Hypothetical | LA1589 |

[1]SPIMP gene numbers are from LA2093. Spimp gene numbers are from LA1589. Solyc gene numbers are from Heinz.
No candidate genes were identified for qRpp1-5.

To further determine if LA1589 resistance to *Pst*19 is Pto/Prf-independent, we analyzed our IBL and F2 mapping data for associations with *Pto/Prf*. The *Pto/Prf* locus is located around the middle of chromosome 5 (Table S6). The delineated regions from the IBL analysis are at the tip of chromosome 5 and do not overlap (Table 1, Figure 4). This is consistent with our finding that *Pst*19 is not recognized by Pto/Prf (Figure 2B), and previous findings that AvrPtoB in *Pst*19 is not recognized (Bao et al. 2015). We also examined whether the delineated regions include *FLS2.1*, *FLS2.2* or *FLS3* since LA1589 has been previously shown to be responsive to flgII-28 from *Pst*19 (Hind et al., 2016). *FLS2.1* and *FLS2.2* are found on chromosome 2 and *FLS3* is found on chromosome 4 (Table S6). We did not find any associations with these two chromosomes and the resistance phenotype (Table 1, Figure 4).

**Discussion**

Genetic resistance is extremely effective in protecting plants from infection. It has been very challenging to identify sources of resistance in wild populations and to have sufficient resolution to identify genes associated with resistance. Here, using our high-throughput seedling flood assay, we identified resistance to *Pst*19 in the wild tomato species *S. pimpinellifolium* LA1589 (Figure 1). Interestingly, although *S. pimpinellifolium* was the original source of the

*Pto/Prf* gene cluster (Pilowsky and Zutra 1982; Pitblado and Macneill 1983; Pedley and Martin 2003), resistance to *Pst*19 in LA1589 is independent of *Pto/Prf*, and does not result in an HR (Figure 2, Figure 3, Table S6). By screening an existing set of IBLs, we were able to narrow the genomic regions of interest to ~250 genes on chromosome 5, ~590 genes on chromosome 6 and ~1335 or ~1025 genes on chromosomes 8, and ~1150 genes in chromosome 12, representing 5 QTL associated with resistance (Table 1, 2, Figure 4). We generated an independent F2 mapping population between Heinz BG-1706 and LA1589 and carried out next-generation sequencing of pools of seedlings with phenotypes at the extremes of resistance or susceptibility. We analyzed variants to identify those that were specific to the resistant pool and missing from the susceptible pool or the susceptible parent cultivar. This analysis enabled us to shortlist eleven candidates as having unique variants in the resistant pool (Table 4).

Further characterization of resistance in *S. pimpinellifolium* LA1589 supports the existence of non-canonical resistance to *Pst*19. LA1589 seedlings supported about 1 log less bacterial growth than MM-*PtoS*, and LA1589 adults supported about 1.3 log less bacterial growth than MM-*PtoS* (Figure 1B, 1D). This difference is much smaller than typical 2-3 log difference observed when ETI is involved (Hassan et al. 2017). Importantly, we see similar resistance in both seedlings and adults, as we have previously observed in other accessions (Hassan et al. 2017). In addition, we did not observe an HR (Figure 3), which is commonly observed with ETI. While *S. pimpinellifolium* LA1589 carries the *Pto/Prf* locus that recognizes AvrPto and AvrPtoB (Figure 2A), it is not sufficient to confer resistance to *Pst*19 (Figure 2B). Consistent with this, the *Pto/Prf* locus is not located in the regions identified for *Pst*19 resistance in either mapping approach (Table 1, 2, Figure 4, Table S6). In addition, previous work showed that AvrPtoB from *Pst*19 is not recognized in *S. habrochaites* (Bao et al. 2015), further supporting that the observed resistance is independent of Pto/Prf. We also confirmed that *FLS2.1*, *FLS2.2* and *FLS3* are not present in any of the delineated regions (Table 1, 2, Figure 4, Table S6). These data further support the presence of additional sources of *Pst*19 resistance in *S. pimpinellifolium* LA1589. Since *S. pimpinellifolium* was introgressed into many early tomato cultivars (Pilowsky and Zutra 1982; Pitblado and Macneill 1983; Pedley and Martin 2003), some resistance to *Pst*19 might already exist in breeding materials (Menda et al. 2014), which would expedite the introduction of new sources of resistance.

Our IBL screen identified specific chromosomal regions that were associated with reduced symptoms and lower bacterial growth (Table 2, Figure 4, 5, 6). These data demonstrate that multiple genes contribute to resistance to *Pst*19 (Figure 4, 5, 6), which is consistent with QDR. No IBLs were available that were homozygous for all resistance QTL identified here. IBLs that carry some but not all linked markers show more disease and smaller reductions in bacterial growth compared to the LA1589 parent (Figure 5, 6). It is important to note that the IBLs have only been genotyped at specific markers. Thus, even though IBLs may be homozygous for the same marker, they may contain different genomic regions from *S. pimpinellifolium* outside the marker. IBLs lacking all linked resistance markers (ie. LA4152) are as susceptible as the cultivar E-6203, with severe disease symptoms and high levels of bacterial growth (Figure 5, 6). This supports the contribution of the QTL to disease resistance. Interestingly, *S. pimpinellifolium* LA1589 was previously identified as having resistance to the *P. syringae* race 1 strain A9 (Thapa et al. 2015). However, the observed resistance to *Pst*A9 was much weaker than the resistance we observed against *Pst*19 (Fig 1), and was not pursued. QDR is a powerful tool in protecting plants from infection, as it typically involves many genes of small to moderate effect. As a result, resistance mediated by QTL is typically harder for

pathogens to overcome (Poland et al. 2009; French et al. 2016). QDR can involve genes with predicted roles in immunity, such as receptor-like kinases and NLRs, as well as genes with different molecular functions (Bao et al. 2015; Debieu et al. 2016; Gonzalez et al. 2017; Guo et al. 2020). Some QTL are associated with effector recognition, such as those involved in the recognition of HopAM1 (Iakovidis et al. 2016) or HopQ1-1 (Luo et al. 2017).

To gain greater resolution into the genes involved in resistance to *Pst*19 within the intervals defined by the IBL analysis, we coupled the IBL screen with a genotyping-by-sequencing screen on a separate F2 mapping population derived from *S. pimpinellifolium* LA1589, and Heinz, which is susceptible to *Pst*19 (Figure S1A). High-quality genome sequences are available for both Heinz and *S. pimpinellifolium* LA1589 (Sato et al. 2012; Wang et al. 2020; Takei et al. 2021). We screened ~1200 individuals for their resistance and susceptibility phenotypes, which did not segregate in a Mendelian manner (Table 3), consistent with the IBL analysis. This population is likely to have more recombination events than the IBLs, which can help identify genes of interest. By focusing on the regions identified by the IBLs and using custom bioinformatic pipelines, we were able to identify unique SNPs that were specific to the resistant pool, but not present in Heinz, the susceptible pool or a susceptible *S. pimpinellifolium* accession (Figure 7). For SNPs in Heinz genes, we identified *S. pimpinellifolium* homologs with high similarity by reciprocal best hit Blast. The *S. pimpinellifolium* genes in the strongly resistant pool had unique differences compared to Heinz or the highly susceptible pool.

Some of our candidate genes have predicted functions that have been shown to be involved in resistance (Table 4). Sumoylation contributes to disease progression and virulence activities of phytopathogenic bacteria, as well as immune responses against pathogens (Park et al. 2011). Mutants in Arabidopsis the SUMO E3 ligase, SIZ1, or SUMO genes, SUM1/2, have high levels of salicylic acid and constitutive expression of pathogenesis-related (PR genes) (Lee et al. 2007; van den Burg et al. 2010). COG2 is one of the subunits of the COG complex, which is believed to anchor vesicles to the Golgi before fusion (Ungar et al. 2006). COG proteins have been implicated in resistance to powdery mildew and soybean cyst nematodes, and are induced in a tissue-specific manner during immune responses (Ostertag et al. 2013; Lawaju et al. 2020; Klink et al. 2021). The AT-rich interactive domain-containing protein 5-like was identified as interacting with the C4 protein, believed to be involved in symptom development, from Tomato Yellow Leaf Curl Virus (Kim et al. 2016). The molecular function of this protein is unknown. Some PP2Cs can regulate PTI and ETI responses to *P. syringae* (Widjaja et al. 2010; Giska and Martin 2019; Sobol et al. 2022). A citrus pectin acetylesterase was found to be a negative regulator of defenses against *Xanthomonas citri* pv. *citri* (Li et al. 2020). Copper chaperone for superoxide dismutase proteins have been implicated in resistance against *Magnaporthe orzyae*, which causes rice blast, and in production of reactive oxygen species, which is associated with immune responses (Li et al. 2019). NADH-ubiquinone oxidoreductase chain proteins are found in complex I of the electron transport chain in mitochondria (Moller 2001). They function in reactive oxygen species production (Moller 2001). Zinc finger CCCH-domain proteins have previously been implicated in immune responses to the fungus *Verticillium nonalfalfae* and *Xanthomonas oryzae* (Svara et al. 2019; Wang et al. 2022). It is also possible that some of these genes are susceptibility factors which are targeted by the pathogen for enhanced virulence (Dangl et al. 2013; Schreiber et al. 2021; Schreiber and Lewis 2021). When susceptibility factors are disrupted, they can enhance plant resistance (Dangl et al. 2013; Schreiber et al. 2021; Schreiber and Lewis 2021).

A previous screen identified QTL on chromosomes 2 and 8 in *S. habrochaites* that contributed to the resistance against *Pst*T1 (Bao et al. 2015), which is very similar to *Pst*19 (Kunkeaw et al. 2010). The QTL on chromosome 8 was not followed up on, however the QTL on chromosome 2 was narrowed to ~140 genes and one candidate gene is a receptor-like protein kinase (Bao et al. 2015). This RLK is distinct from the candidate genes we identified. Another study identified 3 QTL on chromosomes 1, 2 and 12 in *S. habrochaites* that contributed to resistance against a different race 1 strain, *Pst*A9 (Thapa et al. 2015). No candidate genes were identified for these QTL.

The identification of new sources of disease resistance can help bolster plant resilience to infection. Combining classical genetics with next-generation sequencing and high-throughput seedling assays allowed us to identify genomic regions and candidate genes associated with resistance in wild tomato. These data may be helpful for plant breeders in prioritizing loci for introgression and/or for replacement in susceptible varieties.

**References**


Almeida, N. F., et al. (2009) "A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000." Molecular Plant-Microbe Interactions **22**(1): 52-62.

Alonge, M., et al. (2020) "Major impacts of widespread structural variation on gene expression and crop improvement in tomato." Cell **182**(1): 145-+.

Andrews, S. (2010). "FastQC: a quality control tool for high throughput sequence data."

Bao, Z. L., et al. (2015) "Identification of a candidate gene in *Solanum habrochaites* for resistance to a race 1 strain of *Pseudomonas syringae* pv. *tomato*." Plant Genome **8**(3).

Block, A. and Alfano, J. R. (2011) "Plant targets for *Pseudomonas syringae* type III effectors: virulence targets or guarded decoys?" Current Opinion in Microbiology **14**(1): 39-46.

Boyd, L. A., et al. (2013) "Plant-pathogen interactions: disease resistance in modern agriculture." Trends in Genetics **29**(4): 233-240.

Cai, R., et al. (2011) "The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity." PLOS Pathogens **7**(8): e1002130.

Cingolani, P., et al. (2012) "Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift." Frontiers in Genetics **3**(35): 1-9.

Cingolani, P., et al. (2012) "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w(1118); iso-2; iso-3." Fly **6**(2): 80-92.

Couto, D. and Zipfel, C. (2016) "Regulation of pattern recognition receptor signalling in plants." Nature Reviews Immunology **16**(9): 537-552.

Dangl, J. L., et al. (2013) "Pivoting the plant immune system from dissection to deployment." Science **341**(6147): 746-751.

Debieu, M., et al. (2016) "Quantitative disease resistance to the bacterial pathogen *Xanthomonas campestris* involves an Arabidopsis immune receptor pair and a gene of unknown function." Molecular Plant Pathology **17**(4): 510-520.

DeFalco, T. A. and Zipfel, C. (2021) "Molecular mechanisms of early plant pattern-triggered immune signaling." Molecular Cell **81**(17): 3449-3467.

DeYoung, B. J. and Innes, R. W. (2006) "Plant NBS-LRR proteins in pathogen sensing and host defense." Nature Immunology **7**(12): 1243-1249.

Doganlar, S., et al. (2002) "Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589)." Genome **45**(6): 1189-1202.

French, E., et al. (2016) "Mechanisms of quantitative disease resistance in plants." Seminars in Cell & Developmental Biology **56**: 201-208.

Giska, F. and Martin, G. B. (2019) "PP2C phosphatase Pic1 negatively regulates the phosphorylation status of Pti1b kinase, a regulator of flagellin-triggered immunity in tomato." Biochemical Journal **476**: 1621-1635.

Gonzalez, A. M., et al. (2017) "Dissection of resistance genes to *Pseudomonas syringae* pv. *phaseolicola* in UI3 common bean cultivar." International Journal of Molecular Sciences **18**(12).

Guo, Z. F., et al. (2020) "Complex genetic system involved in *Fusarium* ear rot resistance in maize as revealed by GWAS, bulked sample analysis, and genomic prediction." Plant Disease **104**(6): 1725-1735.

Hassan, J. A., et al. (2020) "High-throughput identification of resistance to *Pseudomonas syringae* pv. *tomato* in tomato using seedling flood assay." Jove-Journal of Visualized Experiments(157): 12.

Hassan, J. A., et al. (2017) "A rapid seedling resistance assay identifies wild tomato lines that are resistant to *Pseudomonas syringae* pv. *tomato* race 1." Molecular Plant-Microbe Interactions **30**(9): 701-709.

Heath, M. C. (2000) "Hypersensitive response-related death." Plant Molecular Biology **44**(3): 321-334.

Hirano, S. S. and Upper, C. D. (2000) "Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae* - a pathogen, ice nucleus, and epiphyte." Microbiology and Molecular Biology Reviews **64**(3): 624-653.

Iakovidis, M., et al. (2016) "Effector-triggered immune response in *Arabidopsis thaliana* is a quantitative trait." Genetics **204**(1): 337-+.

Jones, J. D. G., et al. (2016) "Intracellular innate immune surveillance devices in plants and animals." Science **354**(6316): 1117-+.

Kerr, E. A. and Cook, F. I. (1983) "Ontario-7710 - A tomato breeding line with resistance to bacterial speck, *Pseudomonas syringae* pv. *tomato* (Okabe)." Canadian Journal of Plant Science **63**(4): 1107-1109.

Kim, N., et al. (2016) "Comparative analyses of Tomato Yellow Leaf Curl Virus C4 protein-interacting host proteins in healthy and infected tomato tissues." Plant Pathology Journal **32**(5): 377-387.

Klink, V. P., et al. (2021) "Conserved oligomeric Golgi (COG) complex genes functioning in defense are expressed in root cells undergoing a defense response to a pathogenic infection and exhibit regulation by MAPKs." Plos One **16**(8): 30.

Kunkeaw, S., et al. (2010) "Molecular and evolutionary analyses of *Pseudomonas syringae* pv. *tomato* race 1." Molecular Plant-Microbe Interactions **23**(4): 415-424.

Lawaju, B. R., et al. (2020) "The *Glycine max* Conserved Oligomeric Golgi (COG) Complex functions during a defense response to *Heterodera glycines*." Frontiers in Plant Science **11**: 17.

Lee, J., et al. (2007) "Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase." Plant Journal **49**(1): 79-90.

Lewis, J. D., et al. (2009) "The targeting of plant cellular systems by injected type III effector proteins." Seminars in Cell & Developmental Biology **20**: 1055-1063.

Li, H. (2013) "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM." arXiv: 1-3.

Li, H. (2021) "Wgsim." Github repository.

Li, H. and Durbin, R. (2009) "Fast and accurate short read alignment with Burrows-Wheeler transform." Bioinformatics **25**(14): 1754-1760.

Li, H., et al. (2009) "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Li, Q., et al. (2020) "Systematic analysis and functional validation of citrus pectin acetylesterases (CsPAEs) reveals that CsPAE2 negatively regulates citrus bacterial canker development." International Journal of Molecular Sciences **21**(24): 14.

Li, Y., et al. (2019) "Osa-miR398b boosts $H_2O_2$ production and rice blast disease-resistance via multiple superoxide dismutases." New Phytologist **222**(3): 1507-1522.

Lin, N. C., et al. (2006) "Diverse AvrPtoB homologs from several *Pseudomonas syringae* pathovars elicit Pto-dependent resistance and have similar virulence activities." Applied and Environmental Microbiology **72**(1): 702-712.

Luo, Q., et al. (2017) "Genetic interaction between Arabidopsis *Qpm3.1* locus and bacterial effector gene *hopW1-1* underlies natural variation in quantitative disease resistance to *Pseudomonas* infection." Frontiers in Plant Science **8**.

Martin, G. B., et al. (1993) "Map-based cloning of a protein kinase gene conferring disease resistance in tomato." Science **262**(5138): 1432-1436.

Martin, M. (2011) "Cutadapt removes adapter sequences from high-throughput sequencing reads." EMBnet.journal **17**(1): 10-12.

Mazo-Molina, C., et al. (2020) "Ptr1 evolved convergently with RPS2 and Mr5 to mediate recognition of AvrRpt2 in diverse Solanaceous species." Plant Journal **103**(4): 1433-1445.

Mazo-Molina, C., et al. (2019) "The *Ptr1* locus of *Solanum lycopersicoides* confers resistance to race 1 strains of *Pseudomonas syringae* pv. *tomato* and to *Ralstonia pseudosolanacearum* by recognizing the type III effectors AvrRpt2 and RipBN." Molecular Plant-Microbe Interactions **32**(8): 949-960.

Menda, N., et al. (2014) "Analysis of wild-species introgressions in tomato inbreds uncovers ancestral origins." Bmc Plant Biology **14**: 16.

Moller, I. M. (2001) "Plant mitochondria and oxidative stress: Electron transport, NADPH turnover, and metabolism of reactive oxygen species." Annual Review of Plant Physiology and Plant Molecular Biology **52**: 561-591.

Ostertag, M., et al. (2013) "The conserved oligomeric Golgi complex is involved in penetration resistance of barley to the barley powdery mildew fungus." Molecular Plant Pathology **14**(3): 230-240.

Park, H. J., et al. (2011) "SUMO and SUMOylation in Plants." Molecules and Cells **32**(4): 305-316.

Pearson, W. R. (2013) "An introduction to sequence similarity ("homology") searching." Curr Protoc Bioinformatics **Chapter 3**: 3.1.1-3.1.8.

Pedley, K. F. and Martin, G. B. (2003) "Molecular basis of Pto-mediated resistance to bacterial speck disease in tomato." Annual Review of Phytopathology **41**: 215-243.

Pilowsky, M. and Zutra, D. (1982) "Screening wild tomatoes for resistance to bacterial speck pathogen (*Pseudomonas* tomato)." Plant Disease **66**(1): 46-47.

Pitblado, R. E. and Macneill, B. H. (1983) "Genetic basis of resistance to *Pseudomonas syringae* pv. *tomato* in field tomatoes." Canadian Journal of Plant Pathology **5**(4): 251-255.

Poland, J. A., et al. (2009) "Shades of gray: the world of quantitative disease resistance." Trends in Plant Science **14**(1): 21-29.

Ronald, P. C., et al. (1992) "The cloned avirulence gene *AvrPto* induces disease resistance in tomato cultivars containing the *Pto* resistance gene." Journal of Bacteriology **174**(5): 1604-1611.

Salmeron, J. M., et al. (1996) "Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster." Cell **86**(1): 123-133.

Sato, S., et al. (2012) "The tomato genome sequence provides insights into fleshy fruit evolution." Nature **485**(7400): 635-641.

Schreiber, K. J., et al. (2016) "Die another day: molecular mechanisms of effector-triggered immunity elicited by type III secreted effector proteins." Seminars in Cell & Developmental Biology **56**: 124-133.

Schreiber, K. J., et al. (2021) "What the wild things do: mechanisms of plant host manipulation by bacterial type III secreted effector proteins." Microorganisms **9**(5): 48.

Schreiber, K. J., et al. (2021) "Arabidopsis Abscisic Acid Repressor 1 is a susceptibility hub that interacts with multiple *Pseudomonas syringae* effectors." Plant Journal **105**(5): 1274-1292.

Schreiber, K. J. and Lewis, J. D. (2021) "Identification of a putative DNA-binding protein in Arabidopsis that acts as a susceptibility hub and interacts with multiple *Pseudomonas syringae* effectors." Molecular Plant-Microbe Interactions **34**(4): 410-425.

Sobol, G., et al. (2022) "The emerging role of PP2C phosphatases in tomato immunity." Molecular Plant-Microbe Interactions **35**(9): 737-747.

Su, X., et al. (2021) "A high-continuity and annotated tomato reference genome." BMC Genomics **22**(1): 12.

Svara, A., et al. (2019) "Temporal and spatial assessment of defence responses in resistant and susceptible hop cultivars during infection with *Verticillium nonalfalfae*." Journal of Plant Physiology **240**: 12.

Takei, H., et al. (2021) "De novo genome assembly of two tomato ancestors, *Solanum pimpinellifolium* and *Solanum lycopersicum* var. *cerasiforme*, by long-read sequencing." DNA Research **28**(1): 9.

Tang, X. Y., et al. (1996) "Initiation of plant disease resistance by physical interaction of AvrPto and Pto kinase." Science **274**(5295): 2060-2063.

Thapa, S. P., et al. (2015) "Identification of QTLs controlling resistance to *Pseudomonas syringae* pv. *tomato* race 1 strains from the wild tomato, *Solanum habrochaites* LA1777." Theoretical and Applied Genetics **128**(4): 681-692.

Ungar, D., et al. (2006) "Retrograde transport on the COG railway." Trends in Cell Biology **16**(2): 113-120.

van den Burg, H. A., et al. (2010) "Arabidopsis small ubiquitin-like modifier paralogs have distinct functions in development and defense." Plant Cell **22**(6): 1998-2016.

Wang, L. H., et al. (2022) "OsMAPK6 phosphorylates a zinc finger protein OsLIC to promote downstream OsWRKY30 for rice resistance to bacterial blight and leaf streak." Journal of Integrative Plant Biology **64**(5): 1116-1130.

Wang, X., et al. (2020) "Genome of *Solanum pimpinellifolium* provides insights into structural variants during tomato breeding." Nature Communications **11**(1): 11.

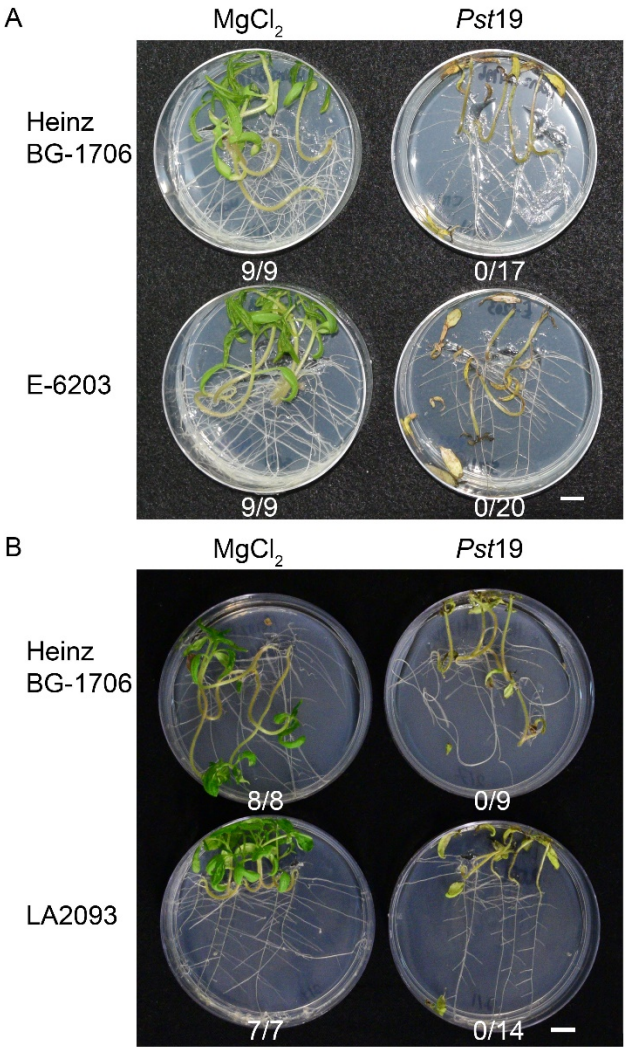
Ward, N. and Moreno-Hagelsieb, G. (2014) "Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: how much do we miss?" Plos One **9**(7): 6.

Widjaja, I., et al. (2010) "A protein phosphatase 2C, responsive to the bacterial effector AvrRpm1 but not to the AvrB effector, regulates defense responses in Arabidopsis." Plant Journal **61**(2): 249-258.

Xin, X. F., et al. (2018) "*Pseudomonas syringae*: what it takes to be a pathogen." Nature Reviews Microbiology **16**(5): 316-328.

Yunis, H., et al. (1980) "Two sources of resistance to bacterial speck of tomato caused by *Pseudomonas tomato*." Plant Disease **64**: 851-852.

Zamir, D. (2001) "Improving plant breeding with exotic genetic libraries." Nature Reviews Genetics **2**(12): 983-989.




**Supplemental Figure S1.** Tomato cultivars Heinz BG-1706 and E-6203 (A) and *S. pimpinellifolium* LA2093 (B) are susceptible to *Pst*19. All seedlings show disease symptoms 10-14 days after being flooded with *Pst*19. The number of surviving plants is indicated over the total number of tested plants.

**Table S1.** Frequency of *Pst*19 resistance in seedlings from recombinant inbred backcross lines

| Inbred Backcross Line | # Susceptible Seedlings | # Resistant Seedlings | Total Seedlings Tested | Overall Resistance (%) |
|---|---|---|---|---|
| LA4139 | 20 | 0 | 20 | 0% |
| LA4140 | 21 | 0 | 21 | 0% |

| | | | | |
|---|---|---|---|---|
| LA4141 | 26 | 0 | 26 | 0% |
| LA4142 | 10 | 0 | 10 | 0% |
| LA4143 | 51 | 2 | 53 | 3.77% |
| LA4144 | 20 | 3 | 23 | 13.04% |
| LA4145 | 32 | 0 | 32 | 0% |
| LA4146 | 20 | 0 | 20 | 0% |
| LA4147 | 27 | 3 | 30 | 10.00% |
| LA4148 | 24 | 3 | 27 | 11.11% |
| LA4149 | 32 | 3 | 35 | 8.57% |
| LA4150 | 19 | 4 | 23 | 17.39% |
| LA4151 | 15 | 0 | 15 | 0% |
| LA4152 | 15 | 0 | 15 | 0% |
| LA4153 | 20 | 0 | 20 | 0% |
| LA4154 | 18 | 0 | 18 | 0% |
| LA4155 | 15 | 0 | 15 | 0% |
| LA4156 | 22 | 14 | 36 | 38.88% |
| LA4157 | 17 | 0 | 17 | 0% |
| LA4158 | 25 | 0 | 25 | 0% |
| LA4159 | 10 | 0 | 10 | 0% |
| LA4160 | 19 | 0 | 19 | 0% |

| | | | | |
|---|---|---|---|---|
| LA4161 | 16 | 0 | 16 | 0% |
| LA4162 | 17 | 0 | 17 | 0% |
| LA4163 | 16 | 0 | 16 | 0% |
| LA4165 | 14 | 0 | 14 | 0% |
| LA4166 | 16 | 0 | 16 | 0% |
| LA4167 | 19 | 2 | 21 | 9.52% |
| LA4168 | 21 | 16 | 37 | 43.24% |
| LA4169 | 15 | 0 | 15 | 0% |
| LA4170 | 25 | 0 | 25 | 0% |
| LA4171 | 12 | 0 | 12 | 0% |
| LA4172 | 24 | 0 | 24 | 0% |
| LA4173 | 24 | 3 | 27 | 11.11% |
| LA4174 | 23 | 0 | 23 | 0% |
| LA4176 | 13 | 0 | 13 | 0% |
| LA4177 | 25 | 0 | 25 | 0% |
| LA4178 | 25 | 0 | 25 | 0% |
| LA4179 | 16 | 0 | 16 | 0% |
| LA4180 | 13 | 0 | 13 | 0% |
| LA4181 | 11 | 0 | 11 | 0% |
| LA4182 | 15 | 0 | 15 | 0% |

| | | | | |
|---|---|---|---|---|
| LA4183 | 9 | 0 | 9 | 0% |
| LA4184 | 10 | 0 | 10 | 0% |
| LA4185 | 14 | 0 | 14 | 0% |
| LA4186 | 8 | 0 | 8 | 0% |
| LA4187 | 19 | 0 | 19 | 0% |
| LA4188 | 10 | 0 | 10 | 0% |
| LA4189 | 17 | 1 | 18 | 5.56% |
| LA4190 | 9 | 0 | 9 | 0% |
| LA4191 | 9 | 0 | 9 | 0% |
| LA4192 | 20 | 1 | 21 | 4.76% |
| LA4193 | 9 | 0 | 9 | 0% |
| LA4194 | 21 | 0 | 21 | 0% |
| LA4195 | 25 | 0 | 25 | 0% |
| LA4196 | 21 | 1 | 22 | 4.55% |
| LA4197 | 25 | 0 | 25 | 0% |
| LA4198 | 10 | 1 | 11 | 9.09% |
| LA4199 | 30 | 2 | 32 | 6.25% |
| LA4200 | 30 | 1 | 31 | 3.23% |
| LA4201 | 8 | 0 | 8 | 0% |
| LA4202 | 15 | 1 | 16 | 6.25% |

| | | | | |
|---|---|---|---|---|
| LA4203 | 24 | 2 | 26 | 7.69% |
| LA4204 | 17 | 2 | 19 | 10.53% |
| LA4205 | 14 | 2 | 16 | 12.5% |
| LA4206 | 10 | 0 | 10 | 0% |
| LA4207 | 17 | 2 | 19 | 10.53% |
| LA4208 | 11 | 8 | 19 | 42.11% |
| LA4210 | 10 | 0 | 10 | 0% |
| LA4211 | 7 | 0 | 7 | 0% |
| LA4212 | 10 | 0 | 10 | 0% |
| LA4213 | 21 | 0 | 21 | 0% |
| LA4214 | 28 | 1 | 29 | 3.45% |
| LA4215 | 6 | 0 | 6 | 0% |
| LA4216 | 24 | 14 | 38 | 36.84% |
| LA4217 | 9 | 0 | 9 | 0% |
| LA4218 | 22 | 0 | 22 | 0% |
| LA4219 | 9 | 0 | 9 | 0% |
| LA4220 | 8 | 0 | 8 | 0% |
| LA4221 | 21 | 0 | 21 | 0% |
| LA4223 | 19 | 0 | 19 | 0% |
| LA4224 | 9 | 0 | 9 | 0% |

| | | | | |
|---|---|---|---|---|
| LA4225 | 2 | 0 | 2 | 0% |
| LA4226 | 24 | 0 | 24 | 0% |
| LA4227 | 35 | 0 | 35 | 0% |
| LA4228 | 37 | 5 | 42 | 11.90% |
| LA4229 | 20 | 1 | 21 | 4.76% |

Inbred backcrossed lines were derived from a cross between E-6203 x *S. pimpinellifolium* LA1589 (Doganlar et al. 2002).

**Table S2.** LA1589[1] and Heinz BG-1706 locations and Wilcoxon Mann-Whitney test for significance of IBL markers used to define chromosomal intervals in LA1589

| Chr # | IBL marker | SGN marker | Significance of marker linked to resistance[2] | Heinz marker coordinates[3] | LA1589 marker coordinates[4] |
|---|---|---|---|---|---|
| 5 | CT101 | SGN-M384 | ** | SL4.0ch05:897017…897148 | Spim0.1ch05: 121112...124083 |
| 5 | TG441 | SGN-M102 | * | SL4.0ch05:1997864…1997884 | Spim0.1ch05: 1159752...1160210 |
| 5 | CT167 | SGN-M188 | ns | SL4.0ch05:2721438…2721899 | Spim0.1ch05: 1915173…1915640 |
| 6 | CT216 | SGN-M516 | ** | SL4.0ch06:730940…730959 | Spim0.1ch06: 702949…703437 |
| 6 | TG178 | SGN-M143 | ns | SL4.0ch06:22112109…22112568 | Spim0.1ch06: 22761978…22762437 |
| 8 | TG176 | SGN-M139 | * | SL4.0ch08:1920661…1921136 | Spim0.1ch08: 1993875...2005857 |
| 8 | TG349 | SGN-M214 | ns | SL4.0ch08:54892017…54892477 | Spim0.1ch08: 55441923...55443207 |
| 8 | TG302 | SGN-M111 | ns | SL4.0ch08:56766398…56766417 | Spim0.1ch08: 57274175...57275344 |
| 8 | TG201 | SGN-M391 | *** | SL4.0ch08:61445780…61446278 | Spim0.1ch08: 62316597...62392315 |
| 8 | CT265 | SGN-M673 | ** | SL4.0ch08:61783768…61784009 | Spim0.1ch08: 62698577…62698818 |
| 8 | CT68 | SGN-M44 | ns | SL4.0ch08:63905392…63905956 | Spim0.1ch08: 64810648...64811212 |
| 12 | TG360 | SGN-M70 | ns | SL4.0ch12:4721141…4721620 | Spim0.1ch12: 4797357...4798344 |
| 12 | TG565 | SGN-M218 | * | SL4.0ch12:58992429…58992903 | Spim0.1ch12: 59447844...59449474 |

| | | | | | |
|---|---|---|---|---|---|
| 12 | TG111 | SGN-M182 | ns | SL4.0ch12:61370469…61370979 | Spim0.1ch12:61434295...61435370 |

[1]LA1589 PacBio genome sequence (Alonge et al. 2020)
[2]Sigificance of markers determined using Wilcoxon Mann-Whitney test with the following p-values. $p>0.1$ not significant (ns); $p<0.1$ suggestive (*); $p<0.05$ significant (**); $p<0.01$ highly significant (***)
[3]From SL4.0_SGNmarker_sequences in ITAG4.0 Release obtained from https://solgenomics.net/ftp/tomato_genome/annotation/ITAG4.0_release/
[4]If marker mapped to more than one region, then region with lowest e-value was selected

**Table S3.** Sequences of RFLP markers

| Chr[1] | IBL Marker | Forward Sequence[2] | Reverse Sequence[2] |
|---|---|---|---|
| 5 | CT101 | CGGAATTCCTTTTTTTTTTTTTTTTTT<br>GAATATGAAATGCTGAAAAAAAATT<br>CAACCTTGCGTGTTCTTGAACTAAAC<br>AATAATCTGATTGACTATTCTGGATT<br>TTCAGGTCTTGCTGGATCACTTCTTG<br>AGAATAAAACTTTACAATCCTTGCAT<br>CTCAACGGCAATTATGGTGGTGCTTT<br>AGGGGCTGCAGCACTGGCTAAAGGC<br>TTGGAGAGCAATAAGTCCTTGAGGG<br>AACTTTATTTGCAGGGGAACTCTGTT<br>GGCGATGAAGGAGTACGGGCATTAA<br>TCTCTGGCCTGTCTTTGCGTAAAGGG<br>AAACTTGTCTTACTTGACCTTGCCAA<br>CAATTCAATTACTGCAAGAGGAGCA<br>TTTCATGTTGCCGAATACGTCAAAAA<br>AAGCAAAAGCTTACTGTGGTTGAAT<br>CTCTATATGAATGACATTAAAGATGA<br>GGGAGCTGAAAAGATTGCAGAAGCT<br>TTGAAGGAGAACCGCTCGATAACTA<br>ATGTTGATCTTG | TTTTTTTTTTTTTTGAATATGAAATG<br>GTGAATAGAAGTAATCTCTATATACT<br>TGCCACTAATGTTTCAGAAACGGAAT<br>GTTATTTATCAAAATTTCCACCGTCC<br>CCCCGCTCAAAATGATTCAGTCACAA<br>AGGGGAGCTGAGGTCGGACTCTTAG<br>GAATATCATTGGTAAAAGAAAAAAA<br>AATACTCACCTATGCAGCCCTGCCCC<br>AACCTCAAGAATGAAAGCAGGGAGA<br>AAAATAAATCTGATCTAAGGAATGT<br>GGTATTTTAAAACATAACAGCAAGCT<br>CCTTCTCAGTCATCTCAAATACATGG<br>TCTCTTGCATCTGTGAGAGCTGTCTG<br>TCCTAATTTGGTGAGGAAGTTACTGG<br>CAAGATTTAAAGATGTGAGTCTAACA<br>TCTTCATTTGCTTTAAGTGCTTGAGC<br>GATAGAAAAAGCTCCTTCATCTCTAA<br>TTTCATTGAACCCTAAATTGAGTGTC<br>ATCAAAGATTCATTGACTACTT |
| 5 | TG441 | TGTATTGAATAATGTCAGCATAGGCT<br>TTTCCAAGCCTGCTCTGAGGTAAACT<br>CTCAGTTCTCTGATTGTCTCCTTGATC<br>CTGAGGCAAAAAGATGTCATACAAC<br>ATTTTGTATCAGTTGGAGGGGGAAG<br>AGGTCAGAGGAGAGGAAGAATGAAC<br>ACACACAACTGAAGGCACAACAAAA<br>TGCCAACATACAATAGATGAAGTTG<br>ATGTGGTGACTCGTTGACCAAGTGG<br>ACTATCCGAAAACATGGTTTCACTTC<br>TTGCATCATATTTTCCTCTGCGAGAG<br>CGTCCCCTCGTAATTATTGCTTTCAG<br>CCCAGTAATCCAAACCTCAGCTTCAT<br>CCTTGTCTTTACAGATCTACACAGAA<br>TTATCATAACAGTTACTAGATTATTT<br>TTTTAAGACTTGTAAATATCAAAATG<br>TCATTTTTCCCGACCGGAAAGAGTAC<br>TTATTTTCATTGAGAAGACAAA | TGGTATGGTACCACGACAAAGAAGA<br>AAAACAGCTTGAACTGTGTCATGTAT<br>CAAGGATTATTCCTGGACAGAGAACT<br>GTGAGTTTGAGTATTCCCACATCTTT<br>CTTCCCTTATAAAGGAAATTTTTTCC<br>CTCAAGTTCCATTTGAATGGAAGATT<br>TGCAGGTTTTACTCTTTAAGAGAAGG<br>TGGATTAAATTTTCTGATCATTGATC<br>AGCAAGAGTAATTTCTGCATCGAGTT<br>CTTTGTTGTGCTCTATGACTGATTTGT<br>AAAGACCATATATTATAGTTCACAAG<br>GGCATTGCTATTCTTATGTCTGTCGA<br>GTTTAGATCCTTTTTCTGTGTGGTTTA<br>TACTCCTCTTCCTCTTGTGCTAGACCT<br>TTATTTTCAAGTGGCTGACCAAAACA<br>GTGGTAATTTTCTCCTCTTCAGGCAA<br>TATTCCAGCGGTATCCTCGACCTGAA<br>AAGGAGTATCAATCATTTTCTC |
| 5 | CT167[3] | ATTAGCATTTGTGCTGGATTGTACTG<br>TTGTTGTTGGCCTCCGTTGTTCATGTT<br>CATTAGCATCGAGGCAGAGGAAGGA<br>TGATGATGACCACCGGTGGATAAAG<br>ATCCCGGTATAGCATTGTTTTGATGA<br>CCTTGAAATCCACCCGCGTTCAAGAA<br>TGCAGCAACGTTGTTGTTATTGTTAG<br>TAGTAGTAGTAGTAGTTCCATTGTTG | Not Available |

| | | | |
|---|---|---|---|
| | | TTATTCCCATGAAAACCAGCAAGATT<br>CATCATTGAGTTTAAGTCGTTTCCTC<br>CCACCATTCCAATTTTCGCAGCCTCG<br>TTGAGGTTGACATTTGGACCTAATTG<br>AGCAGCAGAAGCAGCACCATTCATG<br>TTCATTTTCATAGCATGAATTGCTCT<br>TTGGTCAATGCCACCGCCTCCGGGAT<br>TTGCCTTCATTGGCATATTCGGACTA<br>GGGCCAGTTTTCTTGCCATGAGCAGC<br>AACATTGTTGTTGTTACCTTTC | |
| 6 | CT216 | CTTTTTTTTTTTTTCAAACAACACCAA<br>AAAAAAGTTGAGTATTCACAGCTTA<br>AGTTTTCTTTCAATTGCTATAAAAAA<br>AAACAATGGTCTTTCTATTACTTATA<br>ATTACTGCTACAAATTCAACAAAATG<br>ACACTAAAAATGTATTGAATTTCATA<br>TCAAAGCCTTCTCTATTTACATATTC<br>CCAACAAAATGGCATCTTTTCTCTAT<br>GCCATATAGGCCGCTCTCTTATCCAT<br>CAATCTTCTACTTCTAGTCGACTGCT<br>CTAGAACCAAACTTTCTCTTTTTTTA<br>GTGCTCCCTTCCCACTCAACAGGTAA<br>TAATTTAAAGAACGACATTATTGTCT<br>TCTTCTTCTAATCGACTGCTCTACCA<br>ATCACACATTGCCTTACACTTGGCAT<br>CGGGGGATATCGTCTTCTGCTTCTAG<br>TCGACTGCTCTATAAAAATAATAATT<br>GAATATTGGCGGTTTCCGACTAAGG<br>AATTACCATCTTGTTACTCCATTACG<br>TTCT | TTTTTTTTTCTCTTAGGAAACCTTTTT<br>TATATTCTCCGGCGAGCCAAATCGGA<br>GCATAATTTGGGGGTTTTGAGTTATT<br>GGATCTAACCGGGTTTATTGGGTTGA<br>GTTTTGACTGATCTCCGGCGACTGCA<br>GATGTCTTCGTACTGGTGTTATCGGT<br>GTACAAGGTTTGTTAGAATTTCAGTT<br>GGAAACGATGTTGTTTGTCCTCACTG<br>TGATAGTGGATTTATAGAAGTTGCTG<br>AGGGAAATATTGGGTCGCCGGAATC<br>TCGCCGGAGGTTTCCCATGTGGAACG<br>AACGGCCAGAGTCAGATCGGAGTGT<br>GAACATGGGTTCTCGACGGAGCAGG<br>CGTAATCGGGGTGATAGATCTCCGTT<br>TAATCCGGTTATTGTTCTTCGTAGTCC<br>ATCTGAAGCACCTGGGGAGGATGAA<br>GCTGCTGCAGCTGCAGAGGAGCGGA<br>GTTATGAGCTTTATTATGATGATGGT<br>GAAGGTTCAGGTCTCAGGCCGTTGCC<br>GCCGACGATGTCGGAGTTTTTGATGG<br>GAT |
| 6 | TG178[3] | AATACACCGTTTTCGGCAGGAGAGC<br>AATTCCCCCTACAAACAAAGCCATC<br>AATCATCATGCCTGCACTTGTTGTAA<br>CTGAAGAAACAATGGCCTCCGAGCT<br>ATCCAAGAAAAAAATGAAGAAGACA<br>TCCCAAACAGATATTGAAACTCCTAC<br>TGATAAGAAAACCAAAGAGAAAAAG<br>TCAAAAAAATCCAAGATTGAGTCTG<br>GGTCTGATTCAGAGGACGCTAAAAG<br>GAGTAAAAAGAAAGAGAAAAAGAG<br>AAAAGCTTTGGACTTGGATGGTGAG<br>AAGAGTGATACGAGCTCAGAGATAT<br>GTGAGCCCGTAGATTTGAAGAAGAA<br>TAAAAAGGCTAAATTCGACGACGAA<br>CAAGTGATGGTTGAGAAGAAAGTCG<br>AGGATCCTAATGCTTTGTCTAATTTT<br>AGGATTTCAAAGCCTTTAAAAGAGG<br>CTTTGAATTCTAAGGGAATAGAAGC<br>ACTTTTTCCTATTCA | Not Available |
| 8 | TG176 | AACAGATTATCAAAATGATATACTCT<br>AAAAAAAGAAGATACAAGAGTGTCA<br>ATGGCAGCAAATAAGCTAATGAAAC | TTAAACAAATAAATAGTAGTGGTAGT<br>AATAGCACTGCCCCACATATAATAAC<br>CAAGAGAGAAGAAATTGTAATTAGG |

| | | | |
|---|---|---|---|
| | | TAATTGCTTGTAACAAGTTATGAACA | AGCTCGCATAATCATGTGAGCAGAG |
| | | CAAAACCTGTGCAGGCTGTTGAACC | CTATGAAGTAGGTTAGATATTGGGCT |
| | | GCATTTACACCAGGCATGCCACTTTG | CTGAACTCTGGAGGAATTTGGGGGG |
| | | AGATGTAGATTCTACAGAAAAGACA | ATGTGTTTCTTTGATTTCAACTACAG |
| | | TCAGCCGGTTGACCTGCAATTAGTAG | AAATACATGAAGCCTCGAGTCAATTG |
| | | GCACAGTGTCCGCAAAGGTGACCCT | TCATTTGAAGAAATCTTTTGTTAGAA |
| | | GCTACTAACTGTTGAAGTGCCATTTG | TTAACACCATAGCTTGTGGTACTGTA |
| | | CTTCACCGTTTCTACATAGAACTACA | ATACCTTTTACAAACTAGGAAAGAGT |
| | | ATTATGCAAGAAATCAAAAATATAT | ATTGATGTTCTAGCTGTTACATTTTTG |
| | | TGATAGAAATAACCACACACACACA | ATGAACTTATGTAATTTATTCTCTTCA |
| | | CATGTATAATTGAGACATAGAGAGA | CAATATTTGGCTAAACTTGCCGAAGT |
| | | GAGAGAGAGAGAGAGAGAGAGCGA | TTGCTAAGCTTGCTTTTTTGTGTTCCC |
| | | GAGAGTGAGAGAGGAGGAAGAAGA | TTTTTACCTTGTCAAAAGAAAGTATG |
| | | GGCTGGGTGCCAAAGGGAAACTAAG | TAAAGGTATGTCTTCCTTTTTTTAATG |
| | | AACTAGACAAACCTGATCCCCAAGT | CATTTTAGATTTTTCATTTTGTACAAT |
| | | TGTGCAGCAAGAACAAGAGCTG | CATCTAGCTTGATAGCGTGAAAA |
| 8 | TG349 | TAGGTCAGTACTATTTCTTGAGA | CAGCTATATATGAGAAGTCTAAG |
| | | AACCCCACACACCTGCTGAATTT | CAGGTTTTCTGTTGAAAGTGAAA |
| | | TTTTTTTATTGTAGAATGGTTGAA | CAAAAGTAGAATGGTAAGTATTT |
| | | ATATTGAAATGGATTCACAAGAA | TCTTGCCTCACTTTGAAGAGCACA |
| | | TTTGATTTGTTGGTAAACACAGAT | AGACCATCACCCAAGCTCAAAAG |
| | | AGGAAATCGAGTTAATTGTACAG | CAAAGAAGGAAAAATCAGGAAG |
| | | TCTGTGAAGATTTGGGGAAAAAA | GAAGCAAGAACAATCATAACCAG |
| | | ATGTTCTGTTTACATGATTGCCCT | TCAACCTTGGTCTAAGGACAACC |
| | | TTGTTAATCTGTTGGTTTTTCTTT | AAGATTTGTGTGATATGATGTAA |
| | | ACTTTGGTTATAATAATCCTCTTT | GCTCAAAATGGCAAAAACACAAA |
| | | TGTTTATTGGGGTTTTTGATTGTT | CACAATTCAATAAGACGACGCAT |
| | | TGATTGGAGTGATTGATTCTCCA | ATATTGGCTGTCCACAACTTCCCA |
| | | GAAAGGTCAACGTGTGGGGTGGG | GAAACTAAACTATACAAAATTAT |
| | | GGATTTCGGGGGTNGGGTTTGGT | TTATTGAATTTTCAATATTATTGT |
| | | TGATGACCTANAATTGGGAATTG | CATGACCTTTAAAATCTCAAAACT |
| | | GGGTTCTTGACTTTTGTTAGTGGG | CATTCAATTACTCCACTTATATCA |
| | | GCTATTGCTATTGAATTTTTCTGT | GGGACTATTTGATGATCCATTTAC |
| | | ATGGTACTAGTTTGGTGGATTCTT | CAAGGCATTACCATAAATCTAAT |
| | | ACTGGATTTGTTTATTGGGGATTA | AGCCATCCACAACAAGGCAAAGT |
| | | CTGATTGAAGTAGCT | ATAGAACTAGCTCATAAAACTCC |
| | | | TAATGCTTATCGTTA |
| 8 | TG302 | ATGGTTGCATTACTCAGTCATTGCAC | AAAGAAGCAAGAACCATTTTAATTA |
| | | TCTCCGGGTGGCTATTACATTTTTGA | GGAAAGCATCTAGCTGAAATTAGCG |
| | | AAGGAGTAATTATTTCCCGCAAGAG | GCCAGAGGCACCACCTCTTGCAAGA |
| | | AAGCACTTCCCTTATGTATAGAAATC | AAAGATACTATAACCCTAGCTAGAA |
| | | AGTGACGTAAAATTTATAATCAACAT | AAAGTGCAAATTTTTATGTAAACAGC |
| | | AAAACTCTAAAGTTTGGTTCTGGTGA | ATTGTCCAACAAAAATAAAGGAAGA |
| | | AGTAACTGGAGTTAAATAATCCTACT | CTAAAAACAACATAAATCTAACTTGT |
| | | AACACTCATATCAGAGGGTTGGCTCC | TTGTCAGAACTAGATATCTCAAATAA |
| | | AAACAAATTTCTCGGTCATGAAAAA | ATGAAAGTTTGTCTTATTTCTACTAA |
| | | CACCCATATAAGAGAAAGAATAGAC | TTATCATCTGGCAACTGGCATTTCTA |
| | | ATCCTTATTTTTGGGGAAGGATTTTG | TCCTAGTTGAATAACCTGCAAGTTTA |
| | | GTTCCTATCAGTGAAATTTGGAATCA | TGAGCAAATCTGATTGTTATCTTTTT |
| | | TAGTTATAAGATGAAAAAAATGAGT | ATTTCTTTTTGGTGTAAATTGAATGA |
| | | GGAAAATTACTTATTAAAAAAGAAA | GTAAATGATTAGGAACTCCAATTACA |

| | | | |
|---|---|---|---|
| | | ACAGGGGTGATACCAGCTTAGATGG<br>ATCGATTTAAGTGCTATAAGATGTAG<br>AAGTATTTGCAAGGTGAAAGGACTC<br>TCCCGCGAATTTATTAAAATAAATTT<br>ACATTTATTGGTCACAA | GTTCTTAGAAGCTTTTGCATATGCTA<br>GTCTTGGGACTCCTCCTTTTCTTTCCT<br>TTTTTCTGTTCTCTTTTGGATGTAATG<br>ACAGCACTTTCTTAGTGCTAATTTTG<br>ACATCAATAATATTAACCCTTCATAA<br>A |
| 8 | TG201 | AATACCAAGATTTAGAAAGCTGGTA<br>TCTTTCTTTCTTGCCTACAACAACCTC<br>CGTTAGTTCTCGTGAGGCACCACGTT<br>TGATATATTCCTATCGAAATGTCTTA<br>ACAGGCTTTGCAGCTAAGTTATCAGA<br>AGAAGATATAAAGGAAATGGAAAAA<br>AAGGAAGGATTTGTTTCTGCACGCCC<br>CCAGCAGTTTGTTAGTTTACACACCA<br>CACATAGTGTCAATTTCTTGGGATTG<br>CAACAGAACATGGGGTTCTGGAAAG<br>ACTCTAATTATGGGAAAGGTGTGATC<br>ATCGGAGTTCTAGACACTGGAATTCT<br>TCCTGACCATCCTTCATTTAGTGACG<br>TTGGGATGCCTACCCCACCTGCTAAG<br>TGGAAAGGAGTTTGTGAGTCAAATTT<br>TATGAACAAGTGTAACAAAAAGCTC<br>ATTGGAGCGAGGTCTTACCAACTTGG<br>CAATGGTTCCCCGATAGATGGTAATG<br>GACACGGTACACACACAGCAAGCAC<br>AGCTGCAGGCATG | CTCACTATAGGGAGACAGCCTGCATG<br>CCTACAGCTGTGCTTGCTGTGTGTGT<br>ACCGTGTCCATTACCATCTATCGGGG<br>AACCATTGCCAAGTTGGTAAGACCTC<br>GCTCCAATGAGCTTTTTGTTACACTT<br>GTTCATAAAATTTGACTCACAAACTC<br>CTTTCCACTTAGCAGGTGGGGTAGGC<br>ATCCCAACGTCACTAAATGAAGGAT<br>GGTCAGGAAGAATTCCAGTGTCTAG<br>AACTCCGATGATCACACCTTTCCCAT<br>AATTAGAGTCTTTCCAGAACCCCATG<br>TTCTGTTGCAATCCCAAGAAATTGAC<br>ACTATGTGTGGTGTGTAAACTAACAA<br>ACTGCTGGGGGCGTGCAGAAACAAA<br>TCCTTCCTTTTTTTCCATTTCCTTTAT<br>ATCTTCTTCTGATAACTTAGCTGCAA<br>AGCCTGTTAAGACATTTCGATAGGAA<br>TATATCAAACGTGGTGCCTCACGAGA<br>ACTAACGGAGGTTGTTGTAGGCAAG<br>AAAGAAAGATACCAGCTTTCTAAATC<br>TTGGTATT |
| 8 | CT265[3] | Not available | TTTTTTTTTGCTAGGAAGGGTCAGAG<br>ATTTTCTCTTCCCGATGAGAGGCGGC<br>CTCCTGTGGAGTTCATTGCCTTCTCC<br>AATTGCTTGCTTGGTTGGACCGGGGG<br>TGCCGATGCTTTGACAAGGGGGATTT<br>TTACAGGCCACCTTTTGCCCCTATAA<br>TGCCTCGAGCTACTTTCCTGGAGTAT<br>GGGAATGATCAAGCTTTTTTAGAATT<br>AATTCCCAGTGGTTGAATAGCTGCTG<br>TATTTGTCGAACCTCTCC |
| 8 | CT68 | CAGATCTACAGATGGAAGATCAATT<br>GCAGACGATTGTCCGAGATCACTTGT<br>AGAGCTGGAATCAATGTTACAAGCC<br>AGAAAAGAAGCCTCCTTCAAGCGGG<br>AAAAATCCCTTGCTCATGCTTTTACT<br>CAACAGGAATTGGATGAGATGGATG<br>CTGTTTGTAGTGAAGAAAGAAATGA<br>AAGGGAATTAGAAGAGACAGCGAAT<br>TGGCTAGACGAGTGGATGTCATCAA<br>AGCAATGGAACAGAGGTTCATTTGA<br>CAGAACAGACTCTATAAAGACTGTT<br>GAGAAGGACACGGCTAAGCCATATT<br>GTAACATGGTTCTCAATGCTCTCAAG<br>ACACCACACTCTAGCCCACTTCGCAG<br>ACGGGCTAATAGTCCTCATTATACTG | CAGATCTAGAGATGGAAGCTCAATT<br>GCAGACGATTGTCCGAGATCACTTGT<br>AGAGCTGGAATCAATGTTACAAGCC<br>AGAAAAGAAGCCTCCTTCAAGCGGG<br>AAAAATCCCTTGCTCATGCTTTTACT<br>CAACAGGAATTGGATGAGATGGATA<br>TTGTTTGTAGTGAAGAAAGAAATGA<br>AAGGGAATTAGAAGAGACAGCGAAT<br>TGGCTAGACGAGTGGATGTCATCAA<br>AGCAATGGAACAGAGGTTCATTTGA<br>CAGAAGAGACTCTATAAAGACTGTT<br>GAGATGGACACGGCTAAGCCATATT<br>GTAACATGGTTCCAAATGCTCGAAGA<br>TCACAACACTCTAGCCCACTTCACAG<br>ACAGGCTAGTAGTCCTCATTATACTG |

| | | | |
|---|---|---|---|
| | | CTAATTCTCCCCATCACCAGAGATCA | CTAATTCTCCCCATCACCAGAGATCA |
| | | TCACATTACGATTACTCGGCAATTCA | TCACATTACAATTACTCGGCAATTCA |
| | | ACCACCAGCCACCCCACCCCCTTGTC | ACCACCAGCCACCCCACCCCCTTGTC |
| | | AACCAGAACCTCTTCAAATGC | AACCAAAACCTCTTCAAATGCGCTCA |
| | | | ACAAGCCCACGTAAAAGCCAATCAA |
| | | | CTGCAAACAC |
| 12 | TG360 | AAACCAGATTCTAGTAATAAAATTTT | TAACTCCTGGGAAGATCCACTAAAA |
| | | AACTAGAAAAGAAAAGCAGGATGCG | GAAGACTTGACTCTTCATCAAAAGTA |
| | | GCAAAATGATCTGGGTCCTTTCAATT | TCATTAAGCAATGAGAACTTGAAAAT |
| | | AATATCAAATCTCCACATGTCTATTT | GGAAGCTTTCCGATGTGTATGGTCAA |
| | | TAAGGCTGGGACCCGGACATTTAAT | AAATCGAATCCAAGATTAAGGTTCTT |
| | | GGATAAGCCGATAAGGGGTATATTT | CTTTCCCGATTTTGTTCCTGAAATGTT |
| | | GGTTCAAAGCATAGCAGTAAGGATA | ATAGATATTTTCGTTGATCCAATTCA |
| | | TG | GTGTAAATTAAGTGATGCTAAATTGC |
| | | TATAATAGTGAAACGGAGGATGAGT | TTATTTCTTGATGTTTGAGAATTGAAT |
| | | ATGAAAAATATTTGAAAGTAAAGGG | TTAGGATGTCTTGAGGAGTATAAATG |
| | | GCAAATATGACCCTTTTCTGCAATTA | CCGGCGTGTTTGATGAGATAGGACA |
| | | AAATTGTAAAATACACCATCTTGAA | ATGGGTTCGTGAATCTTTTAATGAAA |
| | | GAAGTTATGTTGCAACTCTTGCTTTA | TTTGTTCGTGCAGAGGGCCTGTTGAT |
| | | ATTATGCTTGAAAGTAGATAGCTATG | CCATCAAAGTCAAGTCTTCCATACCC |
| | | TGCTGACAATATACGACAGTACCTTA | TTTTGTCCATAATGGTGGTTTAGTTA |
| | | ACATATTGATGAAATCTGATAAGAT | AAAAGCTCTTCACTGGTCTGGTCTTC |
| | | AGCCCCAGAACACCTCTCCATATCCT | ACCAGTAAGTTCATTTTTCATTTATTT |
| | | CAATAATTACAACCACTGGTTTCTCA | TAGACGTTTAGTTTGCAAATATGAAA |
| | | TACTTGCCATGGTCACTATACCATGA | CTATGCTGATGGTAACCAAT |
| | | TGCTAAGAGCGAAACATC | |
| 12 | TG565 | ACTAGCATCTCTTGGAGGATGCTGAG | TCAGCAAAATGTCACACAGAGAGTA |
| | | GTGTCAAGTGGTGTTGACCACTCGTT | CAGTAGTAGAGCACAGTAGAGTAGG |
| | | ACCACTGATTCACAGCTGGTGTCTTT | GAGAAGTTGCCTCAAAAGAGGAAAA |
| | | CGAAGCAAGCTTCGTCTGCAAAACA | GAAAAGGTAACGAACCACACATTTG |
| | | AGAATCACACTTTAATCCTCTGTTAC | ACAGCTCAAAACCACTTTACCAATCC |
| | | CTAAAAACAATAGTTGTTTGATGTAA | AAACAAAAAATCATCACATTATCCCT |
| | | TGAAAGAAGAATTTTCACTTCAATGA | CCCTTCTCTCCTTTCTCTATTACTCTC |
| | | TGGAAAGAAAATCTTACAGTTTGAG | ATTTTCCCCAAGTTTCAGGTACCTTTT |
| | | TTTGCTTGCGAAAGTAGCCATTTTCA | TCCTAACATAATCCGCCCATAGTGTT |
| | | TACACCAGTTGAGAAACTTGCTTCTG | CATCATTCAAGATCTGTCCTTTTGAG |
| | | CAATCTATCATTCTCTTCCATTAATA | GAGACTTCATTCCTTACTATGGTCTT |
| | | GCTTGTTCATTGCTGACAGCTTCCTA | CTTTTTTTGATGATTTCTTATGTGAGA |
| | | TTCACACCCTGAAGCCTTGATGACTC | TGTTGAAAACTGGAAAGAAGTGATA |
| | | TTTCCTCTGTTTTTCCCTACATCTATA | AAGATAGGAGGTTTGGTTTCTGGGGT |
| | | CAACTCAAAGAAACAATCAATTATA | TTGTTTATTTTGCTTTACAAGGGTTAA |
| | | CTTCAAATTAATTGGGGTCGCTAAAA | AGATTGGATCTTTTTTAGTTTTGGTAG |
| | | ATGAATCCTTTAGACTAACAACATCC | ATACCCATGTCTAATCTTGTTTCAGA |
| | | CACAAGTCCTTACCCCTACCTCGCAG | ATTCAAAAGGTTGGTACTTTACTGTT |
| | | AGGTAGAGA | TTGCAAGTGGATGACAGAGGAG |
| 12 | TG111 | GGGGGGGGGGTTCGATACACCCTTA | GGGGGGGGAGAAATATTATTAAACT |
| | | GCTCAAGTAAAGCATTGTAAAATCA | TAACACTAGATACATTTTCCTTTTTCT |
| | | AGTAGGAAATGAAATATATTGAAGC | TTTTTGAGTTTCTCCTTTCTATATAAT |
| | | AACAGCTGGGCATACATCACCAAAA | CCACCAAATGCGAAAAAGTATTGGG |
| | | CAAATTGCCAACCCGGACAAAGAAT | GAAGTGATTAAACAGGACATGACAC |
| | | TGTGGTAATTTGGTTTTAATACGGTA | TTGTAGGCTGTAGCTTACTGAGGATA |
| | | TTCAGGAGTAAAATTTCTGAATACGG | TGACCCTAAATAGGAATATATAGGTC |

| | |
|---|---|
| TATTCGGTACAATATACTAATACCAT | GAGGATCAGGGTAAAAGGTTTGATT |
| TCGATATTCACTAGGATTCTTTTGCC | CTTCCACAACTAAATTAGTTCTATAC |
| CACCATTTCTTCCTAGAACACTGAAC | AAGAAGATTAAAAGGGAAGAAAACA |
| AATCACAATAGAATGGTCGGACAAG | AATCATAAAAGAGTATTGAATCTGA |
| CACCATCTAATCTGAGCCTGACTATC | AACTAAAGATTTGAGATAGTCAGGCT |
| TCAAATCTCTAGTTTCAGATTCAATA | CAAATTAAATGGGGCTTGTCCGACCA |
| CTCTTTTATGATTTGTTTTCTTCCCTT | TTCTATTGTGATTGTTCAGGGTTCTA |
| TTAATCTTCTTGTATAGAACTAATTT | GGAAAAAATGGGGGGCAAAAAAATC |
| AGTTGTGGAAGAATCAAACCTTCTAC | CTAGGGAATATCGAATGGTATTAGTA |
| CCTGATCCTCGACCTATATATTCCTA | TATTGTACCGAATACCGTATTCAAAA |
| TCTAGGGTCATATCCTCAGTAAGCTA | ATTTTACTCCTGAATACCGTATTAAA |
| CAGCCTACAAGTGTCATGTCCTGTCT | ACCAAATTACCACAATTCTTTGTCCG |
| AATCACTTCCCCAATACTTCTTCGCA | GGTTGGCAATTTGTTTTGGTGATGTA |
| TTTGGTGG | TGCCCAGCTG |

Forward and reverse sequences were blasted to LA1589 PacBio genome sequence (Alonge et al. 2020) to delineate regions associated with LA1589 resistance. CD40, CT111, TG330, TG505, TG201 and CT265 (all on chr 8) sequences were not available. The next closest marker was used to delimit regions.
[1]Chr indicates chromosome.
[2]https://solgenomics.net
[3]Only one sequence was available.

**Table S4.** Presence of individual homozygous or heterozygous LA1589 markers linked to *Pst*19 resistance in eight Inbred Backcross Lines

| Line # | Markers[1] and chromosome numbers (Chr) | | | | | | |
|---|---|---|---|---|---|---|---|
| | CT101** | CT441* | CT216** | TG176* | TG201*** | CT265** | TG565* |
| | Chr 5 | Chr 5 | Chr 6 | Chr 8 | Chr 8 | Chr 8 | Chr 12 |
| LA4144 | - | - | HM | HM | HM | HM | - |
| LA4156 | - | - | HM | - | - | - | - |
| LA4168 | - | - | - | HM | - | HM | HM |
| LA4173 | - | - | HM | HM | HM | HM | - |
| LA4208 | HM | HM | HM | - | HM | HET | - |
| LA4216 | - | - | HET | HM | - | - | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| LA4148 | - | - | - | - | HM | HM | HM |
| LA4152 | - | - | - | - | - | - | - |

[1]Significance of markers determined using Wilcoxon Mann-Whitney test with the following p-values.

***Highly significant markers ($p<0.01$)

**Significant markers ($p<0.05$)

*Suggestive markers ($p<0.1$)

LA4152 is a susceptible IBL

HM indicates homozygous LA1589

HET indicates heterozygous LA1589/E-6203

- indicates homozygous E-6203

**Table S5.** Metrics for Illumina whole-genome sequencing

| Sample pool[1] | Mapped to | Genome coverage | Total reads after trimming[2] | Total reads passing quality filter | Percentage of reads passing BWA-MEM quality filter |
|---|---|---|---|---|---|
| LA1589 SR | LA1589 | 388X | 1972121919 | 1433302145 | 72.7% |
| LA1589 HS | LA1589 | 349X | 1982856536 | 1286732013 | 64.9% |
| LA1589 SR | Heinz BG-1706 | 397X | 1970656997 | 1466775692 | 74.4% |
| LA1589 HS | Heinz BG-1706 | 358X | 1980465396 | 1319936306 | 66.6% |
| LA1589 SR | LA2093 | 395X | 1970842936 | 1457012516 | 73.9% |
| LA1589 HS | LA2093 | 354X | 1981293878 | 1308252988 | 66.0% |

[1]SR is strongly resistant and HS is highly susceptible.
[2]Reads were trimmed using Cutadapt v2.4 (Martin 2011).

**Table S6.** Loci and coordinates for previously known tomato genes and LA1589 orthologs involved in resistance to *P. syringae*

| **Locus** | **Locus Name** | **Solgenomics locus** | **LA1589 ortholog** | **LA1589 ortholog coordinates** | **LA1589 gene within IBL region?** |
|---|---|---|---|---|---|
| *FLS2.1* | Flagellin Sensing 2 | Solyc02g070890.3.1 | Spim02g021150.1.1 | Spim0.1ch 02: 41246499-41250461 | No |

| | | | | | |
|---|---|---|---|---|---|
| *FLS2.2* | Flagellin Sensing 2 | Solyc02g070910 | Spim02g021180.1.1 | Spim0.1ch02: 41254163-41257963 | No |
| *FLS3* | Flagellin Sensing 3 | Solyc04g009640.2.1 | Spim04g009330.1.1 | Spim0.1ch04: 5202597-5206136 | No |
| *PRF* | Pseudomonas resistance | Solyc05g013280 | Spim05g011870.1.1 | Spim0.1ch05: 5340155-5345742 | No |
| *PTO* | Pto-like serine/threonine kinase | Solyc05g013300.1.1 | Spim05g011900.1.1 | Spim0.1:ch05: 5368138-5369109 | No |

## CONCLUSION

In this work, we identify natural resistance to *P. syringae* in a wild tomato population. This is a key step towards rational design of *P. syringae* resistance in crops.

## CHAPTER VI: Computationally designed protein linkers for ZAR1-guardee nanobody fusions do not result in activation of plant immunity in the presence of antigen

### PREFACE

The work presented in this chapter was a collaboration with Kyungyong Seong, Ksenia V. Krasileva and Jennifer D. Lewis. I performed the transient gene expression and HR assays, co-IP assays and western blots. I performed computational design of linkers in collaboration with Kyungyong Seong and wrote the text with Jennifer D. Lewis. The data presented here has been reproduced with permission of all collaborators.

Incorporating adaptive immunity components into plant innate immunity recognition pathways may enable rapid design of plant resistance to targeted pathogens. ZAR1 is uniquely suited for this approach, given its conservation across plant species. Plant recognition of pathogen effectors through nanobodies would be a key step towards rational design of *P. syringae* resistance.


### Abstract

The highly conserved plant immune receptor, ZAR1, recognizes pathogens by guarding host proteins for pathogen modification. Upon modification, ZAR1-guarded host proteins cause a clash with the ZAR1 nucleotide binding domain, leading to activation of plant immunity. Here, we attempt to mimic the clash leading to ZAR1-guardee activation by fusing a single chain GFP nanobody to the ZAR1 guardee RKS1 and co-expressing the associated GFP antigen. We used Rosetta and RFdiffusion to computationally design protein linkers that fuse RKS1 and the GFP nanobody in the correct orientation relative to ZAR1 nucleotide-binding domain to maximize potential for activation. While existing protein design tools were used in our approach, our application for linker design to achieve specific orientation of fused protein domains has not been extensively explored. We transiently expressed RKS1-GFP nanobody fusions with *A. thaliana* ZAR1 and GFP in *N. benthamiana* and screened for activation of ZAR1. While we were unable to achieve antigen-mediated activation of plant immunity through ZAR1, our computationally designed linkers enabled complex formation between ZAR1, RKS1, and the GFP antigen. In addition, our linker design approaches may have future applications for protein fusions across biology. Future computationally-aided protein design approaches may enable more accurate modeling of custom linkers, and ultimately the success of antigen-mediated ZAR1 activation.


### Introduction

New machine learning techniques may enable designer proteins, with applications across nearly all fields of biology, including plant-microbe interactions. The orientation of protein domains in protein-protein fusions can dictate the function of each domain. Flexible linkers are typically used for protein-protein fusions, as they maximize the functionality of fused protein domains by reducing steric hinderance (Ikebe et al., 1998; Gokhale and Khosla, 2000). However, in certain circumstances it may be necessary for protein domains to be in a specific orientation relative to each other, for example to regulate their activation. Improper orientation may occlude protein-protein interactions, or prevent activity of fused proteins (Maeda et al., 1997; Amet et al., 2009). In these instances, proteins may need to be connected via tailored rigid linkers.

Despite the importance of protein linkers in protein-protein fusions, there have been few attempts to generate tailored rigid linkers computationally. Most linker design efforts have been

directed towards empirically tuning the length, rigidity or flexibility of linkers to achieve proper separation and activity of fused protein domains (Chen et al., 2013). Protein fusion linkers are often separated into flexible or rigid classifications. Flexible linkers are typically glycine-rich, with non-polar serine or threonine amino acids included to increase solubility (Chen et al., 2013). The length of flexible linkers can be adjusted to achieve proper distance between fused proteins to prevent hindrance while permitting a certain degree of movement (Maeda et al., 1997; Amet et al., 2009). When separation of proteins is key for stability or activity, rigid linkers are often advantageous. Rigid linkers typically adopt an α-helix structure, and are often composed of alanine residues together with charged amino acids to form salt bridges in the helical structure (Chen et al., 2013). Multidomain proteins that exist in nature appear to adopt similar linker architecture with either rigid α-helical or flexible linkers, based on natural protein domain linkers with available crystal structures (George and Heringa, 2002). While no tools currently exist specifically for the computational design of linker proteins, general protein design tools enable design of protein backbones and refinement of amino acids within the linker region (Kuhlman et al., 2016; Watson et al., 2023).

There is great interest in rapidly designing novel immune receptors to emerging pathogen threats. However, this is a challenge in plants, as they lack an adaptive immune system, and rely entirely on innate immune receptors for resistance to pathogens (Jones and Dangl, 2006). Incorporating animal adaptive immune system components into plant innate immune systems may enable resistance tailored to any pathogen. Since antibodies can be raised against virtually any targets, plant resistance to new pathogens can be expedited. Recently, incorporating antibodies as fusions with the rice nucleotide-binding leucine-rich repeat (NLRs) receptor Pik-1 resulted in plant-mediated immunity to pathogens expressing the nanobody ligand (Kourelis et al., 2023). The rice NLR Pik-1 has an integrated heavy metal associated domain (HMA) near its N-terminus. By swapping a single-chain antibody (nanobody) with the HMA domain, the authors were able to achieve resistance in the presence of the nanobody ligand (Kourelis et al., 2023). Although this nanobody approach was successful with the rice NLR Pik-1, it was dependent on the non-canonical protein domain architecture of Pik-1 and as a result its feasibility outside of rice is unclear. Therefore, utilizing native NLRs that are functional with downstream signaling components, and activated by a universal guardee is an appealing approach.

The unique and highly conserved angiosperm immune receptor, ZAR1, is an excellent system for rational design of novel plant immunity. ZAR1 orthologs are found in nearly all dicots including important crops (Lewis et al., 2010; Gong et al., 2022; Adachi et al., 2023). ZAR1 functions by guarding host proteins in the ZED1-RELATED KINASE (ZRK) family for pathogen modification, and ZRKs from one species can function with ZAR1 orthologs from another species (Lewis et al., 2013; Baudin et al., 2017; Gong et al., 2022). Thus the design of one ZRK could facilitate resistance across multiple plant species. A cryo-EM structure exists for ZAR1 in several stages of activation, upon recognition of the *Xanthomonas campestris* pv. *campestris* effector AvrAC through the ZRK RKS1. The ZAR1 nucleotide-binding domain (hereafter $ZAR1^{NBD}$) functions as a molecular switch for the activation of immunity. Upon pathogen recognition, the $ZAR1^{NBD}$ is pushed away from the ZAR1-ZRK complex via an allosteric clash with the RKS1 activation loop, leading to conformational changes in ZAR1 and ultimately the assembly of a ZAR1 pentamer, termed the resistosome. The ZAR1 resistosome functions as a non-selective cation pore at the plasma membrane (Bi et al., 2021). Calcium movement through this pore leads to NLR-mediated immunity, which is often associated with a programmed cell death known as the hypersensitive response (Heath, 2000).

We sought to activate the ZAR1$^{NBD}$ molecular switch via GFP binding. We attempted to accomplish this by fusing a GFP nanobody to RKS1 and orienting the nanobody towards ZAR1$^{NBD}$ via a rigid linker. We showed that RKS1-nanobody fusions can interact with both ZAR1 and GFP. While we were unable to achieve ZAR1 activation using current computational design methods, future methods may allow for more accurate design and successful execution of our concept. Our computational approaches for linker design may have applications for other problems in molecular biology. Our cloning method enables rapid testing of new linkers as new computational techniques arise. Furthermore, our methods may facilitate the design of rigid linkers to achieve specific confirmations between two fused proteins.

**Methods**

*Molecular cloning*

The GFP$^{nano}$ sequence was codon optimized for *A. thaliana* and synthesized. We used the generated sequence as a template to amplify the RKS1-BspQI-GFP$^{nano}$ fragment via overlap extension PCR. This fragment was amplified to include attB sites and cloned into pDonr207 using the Gateway BP reaction. The final 6 amino acids of RKS1 were excluded, as these amino acids are not well resolved in the ZAR1 inactive structure and would result in the nanobody fusion being further from ZAR1$^{NBD}$. To insert linkers into the RKS1-BspQI-GFP$^{nano}$ vector we used a Golden Gate assembly approach. Complimentary oligonucleotides were synthesized to contain sticky ends matching the vector after it was digested with BspQI. When more than one primer set was needed to generate a linker, additional overlaps were added so that the two primers sets would anneal, and the two primers sets were ligated in the same reaction with the digested vector. After linkers were ligated into the RKS1-BspQI-GFP$^{nano}$ vector, a Gateway LR reaction was performed to transfer the gene fusions into the previously described Gateway compatible pMac16 vector (Diplock et al., 2023), which was modified from pBD to include a C-terminal 3xFlag epitope tag. To generate dexamethasone-inducible GFP, we amplified GFP and inserted it into the previously described pMac15 vector (Baudin et al., 2017) between the XhoI and SpeI sites via restriction enzyme digest and ligation, resulting in a loss of the 5xMyc tag present in the pMac15 vector between the StuI and SpeI sites. The AtZAR1-5xMyc and AtZAR1$^{LRR}$-5xMyc constructs were previously described (Lewis et al., 2013). Clones in pDonr207 was confirmed by sequencing, and clones from the Gateway LR reaction were confirmed by colony PCR.

*Plant growth conditions, VIGS, transient expression, and HR assays*

*N. benthamiana* growth conditions have been previously described (Baudin et al., 2017). Briefly, *N. benthamiana* seeds were sterilized in 10% bleach, germinated on Sunshine #1 (Sun Grow Horticiulture), transplanted to Sunshine #1 supplemented with 14:14:14 Osmocote, and grown at 22°C under 9 h days.

Transient gene expression and VIGS targeting *NbZAR1* were performed as previously described (Baudin et al., 2017). *A. tumefaciens* was grown overnight at 28°C in LB with appropriate antibiotics. Bacteria were pelleted and resuspended in induction media (10 mM MES pH 5.6, 200 μM acetyosyringone), and incubated in the dark for 3-4 h. The optical density of each culture was adjusted to an $OD_{600}$ of 0.25. For VIGS targeting, the fourth and fifth leaves of 3-week-old *N. benthamiana* plants were infiltrated completely with a needle-less syringe. For transient expression, inoculum was infiltrated into leaves of 5-7 week *NbZAR1*-VIGS plants, using a needle-less syringe. For transient gene expression, dexamethasone (40 μM) was sprayed

on leaves 24 h after induction with *A. tumefaciens*. The HR was scored 24,and 48 h after dexamethasone application based on a previously described 4-point scale (Diplock et al., 2023).

*Linker design and protein modeling*

A custom Rosetta script was used to design linkers modified from a previously described method (Kuhlman et al., 2016). A RKS1 and $GFP^{nano}$ heterodimer protein structure input for Rosetta was generated using Phyre2 (Kelley et al., 2015), by modeling the complete sequence of RKS1 and $GFP^{nano}$ as one protein, without a linker sequence. To amend structural changes in RKS1 caused by low confidence modeling of the RKS1 activation loop, the Phyre2 structure was used to position the unlinked RKS1 and $GFP^{nano}$ cryo-EM structures, and this heterodimer was utilized for subsequent Rosetta linker design. As an initial step, Rosetta Relax was used to refine the Phyre2 heterodimer model. The Rosetta BridgeChains command was used to connect RKS1 with $GFP^{nano}$ with the centroid score-function based on input amino acid length and secondary structure (see Figure 3). Amino acid identities of linker backbone structures were designed using the Rosetta PackRotamersMover which fixes the protein backbone position and optimizes sidechain conformation using a Monte Carlo based assessment, followed by MinMover which refines the backbone position after amino acid identities have been determined (Kuhlman et al., 2016). For each specified amino acid length and secondary structure, 1000 models were generated. All protein modeling was performed in UCSF Chimera (Pettersen et al., 2004).

To generate a RKS1, $GFP^{nano}$ heterodimer model that maximized stability of the two proteins, LZerD was used to dock RKS1 with $GFP^{nano}$ with default constraints (Christoffer et al., 2021). The top 20 models from LZerD were manually evaluated using Chimera (Pettersen et al., 2004) to identify a heterodimer that resulted in the correct orientation of RKS1 and $GFP^{nano}$ relative to $ZAR1^{NBD}$. In addition to the LZeRD heterodimer model, the Rosetta model for the R1 linker was used as one input for RFdiffusion linker design (see Figure 4). Next, RFdiffusion was used to generate protein backbone structures for linkers, specifying a linker length of between 5 and 30 amino acids. RFdiffusion generated 10 backbones for each input across 10 separate runs, for a total of 100 linkers per input. Amino acid identities of the top scoring models were assigned using ProteinMPNN (Wang et al., 2022), and Alphafold (Evans et al., 2022) was used to validate the overall similarity of the protein structure following ProteinMPNN design with the RFdiffusion initial backbone design.

*Co-immunoprecipitation experiments and western blotting*

Co-immunoprecipitation assays for identifying interactions among ZAR1, ZED1 and GFP were based on a previously described method (Diplock et al., 2023). 24 h after dexamethasone induction, 3 $cm^2$ of tissue was collected from 2 separate *N. benthamiana* plants. Tissue samples were ground in liquid nitrogen and homogenized in 1 mL IP1 buffer (50mM HEPES, 50 mM NaCl, 10 mM ethylenediaminetetraacetic acid (EDTA), 0.2% Triton X-100 and 0.1 mg/mL Dextran [Sigma-Aldrich D1037, St. Louis, MO], pH 7.5). Samples were incubated with 5 µL Myc-Trap magnetic beads for 2.5 h on a spinning wheel at 4°C. Beads were washed with four short washes (tubes inverted 6 times) followed by two long washes (5 min on spinning wheel) with IP2 (50 mM HEPES, 250 mM NaCl, 10 mM EDTA and 0.1% Triton X-100, pH 7.5). Both IP1 and IP2 were supplemented with a homemade protease inhibitor cocktail (1 mM PMSF, 1 mg/mL leupeptin [Sigma-Aldrich L2023], 1 mg/mL aprotinin [Sigma-Aldrich A6191], 1 mg/mL antipain [Sigma-Aldrich A1153], 1 mg/mL chymostatin [Sigma-Aldrich C7268] and 1

mg/ml pepstatin [Sigma-Aldrich P5315]). The beads and input fraction were boiled with Laemmli buffer prior to western blotting.

For western blotting, the Myc, Flag tagged proteins, or GFP were separated on a 8%, 10% or 12% SDS-PAGE gel respectively. Proteins were detected using an anti-Myc antibody (Abcam ab32072) anti-Flag antibody (Sigma-Aldrich F1804) or anti-GFP (Takara 632598), followed by a secondary antibody coupled to horseradish peroxidase.

## Results

*Testing empirical and nature-based linkers*

We generated a Gateway donor vector containing RKS1, a GFP-nanobody (hereafter GFP$^{nano}$), and a BspQI restriction site to enable Golden Gate Assembly of linkers via complimentary oligonucleotide primers. Insertion of linkers at the restriction site results in an in-frame RKS1- GFP$^{nano}$ fusion (Figure 1A). We selected RKS1 due to the availability of the ZAR1-RKS1 cryo-EM structure (pdb:6j5w) (Wang et al., 2019a; Wang et al., 2019b), and GFP$^{nano}$ due to the availability of the crystal structure of GFP in complex with GFP$^{nano}$ (pdb:3ogo) (Kubala et al., 2010).

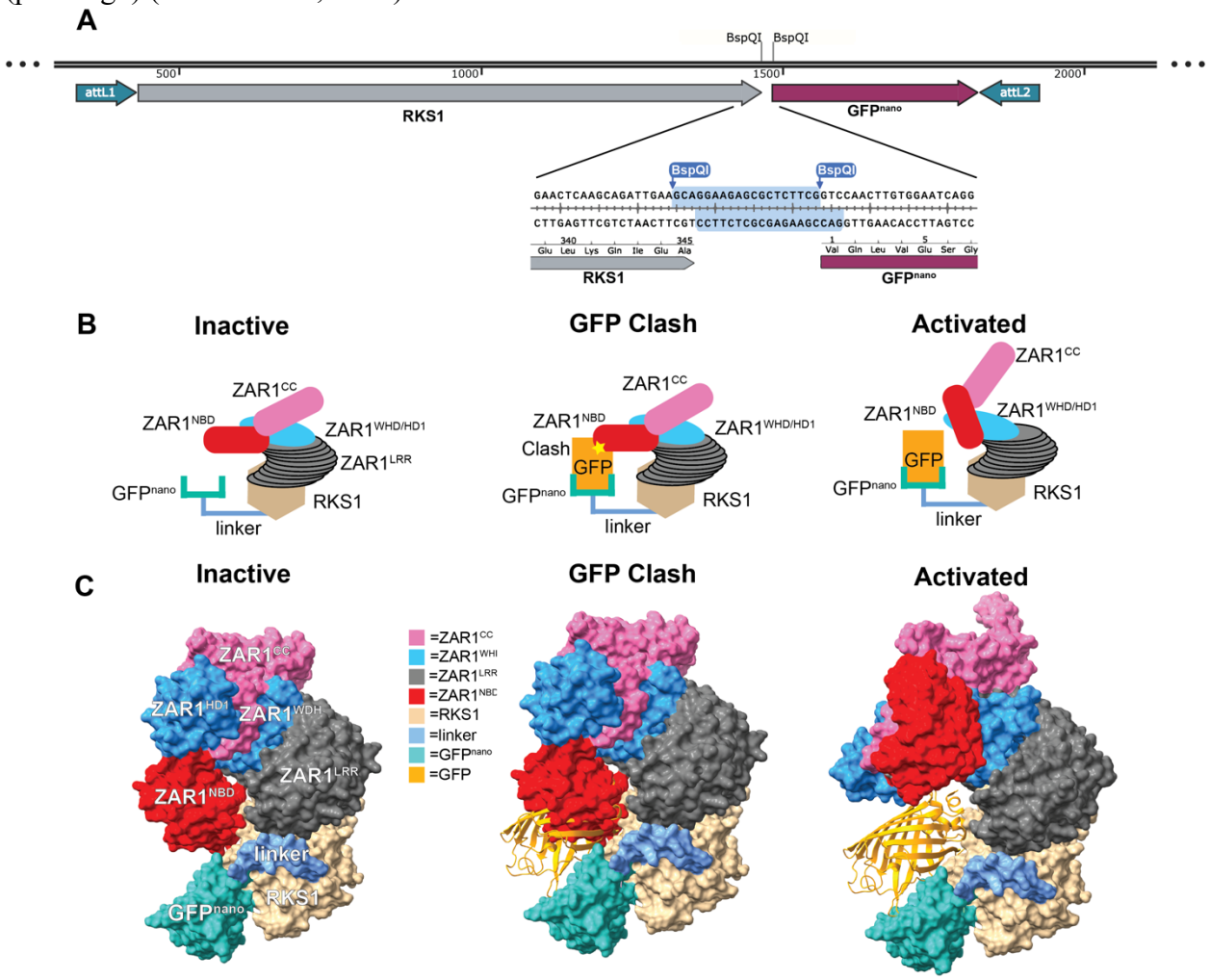



**Figure 1**. The theoretical and structural basis for RKS1-GFP$^{nano}$-mediated activation of ZAR1. (A) diagram of our RKS1-GFP$^{nano}$ in a Gateway-compatible subcloning vector, which enables rapid Golden Gate assembly of linkers using oligonucleotides and *Bsp*QI restriction enzyme digestion. (B) Simplified illustration and (C) structural models

of AtZAR1 in complex with RKS1-GFP$^{nano}$ in inactive conformation, an antigen-induced ZAR1$^{NBD}$ clash conformation, and an activated conformation based on ZAR1-RKS1 structures and GFP$^{nano}$ GFP structures (Kubala et al., 2010; Wang et al., 2019a; Wang et al., 2019b). Structural models were produced in Chimera (Pettersen et al., 2004).

We randomly selected a helical and flexible linker from the protein Linker Database (George and Heringa, 2002), in addition to the commonly used flexible xTen linker, and cloned these into our vector (Table 1). We have previously shown that RKS1 causes an autoactive HR phenotype when transiently expressed in *N. benthamiana* and that this phenotype is dependent on *N. benthamiana* ZAR1 (hereafter NbZAR1) (Baudin et al., 2017; Gong et al., 2022; Diplock et al., 2023). While other *A. thaliana* ZRKs do not result in an autoactive phenotype in *N. benthamiana* and can function in effector recognition with NbZAR1 (Diplock et al., 2023), the structural data of RKS1 led us to utilize this ZRK for nanobody fusions. To overcome RKS1 autoactivity, we tested our constructs in *N. benthamiana* plants silenced for *NbZAR1* (hereafter *NbZAR1*-VIGS). We transiently co-expressed each RKS1-linker-GFP$^{nano}$ construct with *A. thaliana* ZAR1 (hereafter AtZAR1) and GFP. After spraying the plants with dexamethasone, we scored the HR at 24 and 48 h after gene induction. We did not observe any HRs resulting from RKS1-GFP$^{nano}$ fusions based on random linker selection (Figure 2A, B). As expected, the HR was still observed when we transiently co-expressed AtZAR1, ZED1 and HopZ1a (Baudin et al., 2017). *NbZAR1* silencing was effective as we did not observe an HR when ZED1 and HopZ1a were co-expressed. The HR was specific to the co-expression of AtZAR1, ZED1 and HopZ1a, as co-expression of AtZAR1, RKS1 and GFP did not result in an HR.

**Table 1**. Linkers tested and the method used to select linkers

| Linker | Sequence | Length (aa) | Method |
|---|---|---|---|
| L1 | TLRKRKMYEEF | 11 | Random |
| L2 | GGSGGGGG | 8 | Random |
| xTen | GGSGSGAGSGSSGSETPGTSESATPES | 27 | Empirical |
| R1 | DAKDEDLKKTWKKQSSEKGE | 20 | Rosetta |
| R2 | DQTEEARKLLKQAEKDKKKI | 20 | Rosetta |
| R3 | GTTDESKFKETLKRWWKENTNT | 22 | Rosetta |
| R4 | KPDTEEAKKAKKAIKTLFLSLP | 22 | Rosetta |
| D1 | AYDKALGLPVEEMQ | 14 | RFdiffusion |
| D2 | AGTLTPQ | 7 | RFdiffusion |
| D3 | AYALAQGEKPEPGQ | 14 | RFdiffusion |
| D4 | VRGDPEAAKKVLERIKTGINP | 21 | RFdiffusion |

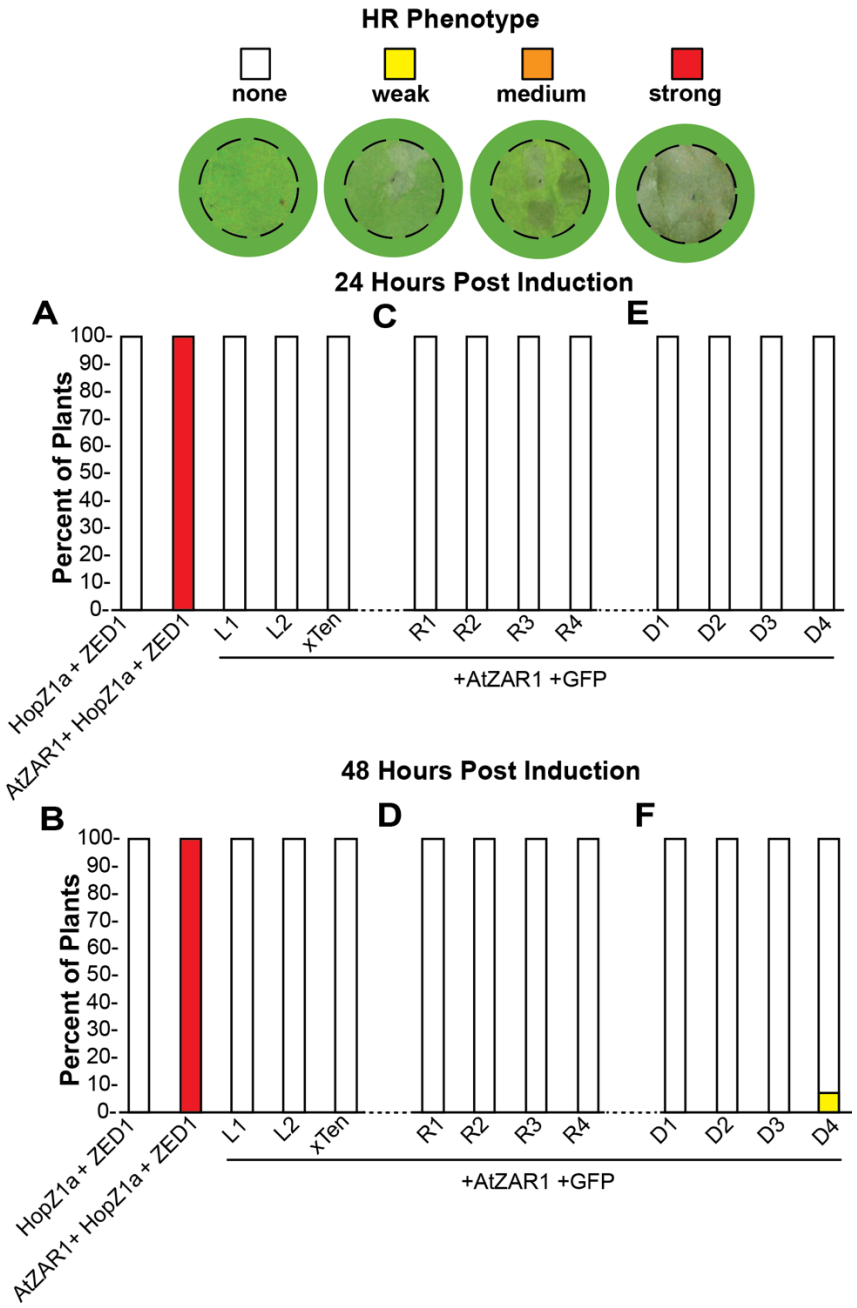



**Figure 2**. RKS1-GFP$^{nano}$ fusions do not cause an HR when co-expressed with AtZAR1 and GFP at two timepoints. RKS1-GFP$^{nano}$ fusions were cloned into a dexamethasone-inducible expression vector and transiently co-expressed with AtZAR1 and GFP in *N. benthamiana* silenced for *NbZAR1* via virus-induced gene silencing (VIGS) (Baudin et al., 2017). The HR was scored on a previously described 4-point scale (Diplock et al., 2023) 24 hours (A, C, E) and 48 hours (B, D, F) after dexamethasone induction (n=14-19 plants). To ensure proper silencing of *NbZAR1* through VIGS, HopZ1a and ZED1 were co-expressed in each *NbZAR1*-VIGS plant (Baudin et al., 2017). As a positive control for the HR, AtZAR1 was co-expressed with ZED1 and HopZ1a (Baudin et al., 2017). Linkers selected using a random approach (A, B), our Rosetta linker design method (C, D) and our RFdiffusion linker design method (E, F) were tested in the same experiment (see Table 1).

*Designing linkers using Rosetta*

To generate rationally designed linkers, we used a custom Rosetta pipeline based on a previous linker design approach (Kuhlman et al., 2016)(Figure 3A). To generate a heterodimer of RKS1 and GFP$^{nano}$ protein structures for Rosetta, we used Phyre2 to generate a complete model of an RKS1- GFP$^{nano}$ fusion without a linker sequence joining the two proteins. The activation loop region of RKS1, which was not resolved in the AtZAR1-RKS1 inactive structure, was modeled with low confidence in the Phyre2 structure, altering other regions of RKS1. To address this, we used the Phyre2 structure to position the unlinked RKS1 and GFP$^{nano}$ cryo-EM structures, and this heterodimer was utilized for subsequent Rosetta linker design.

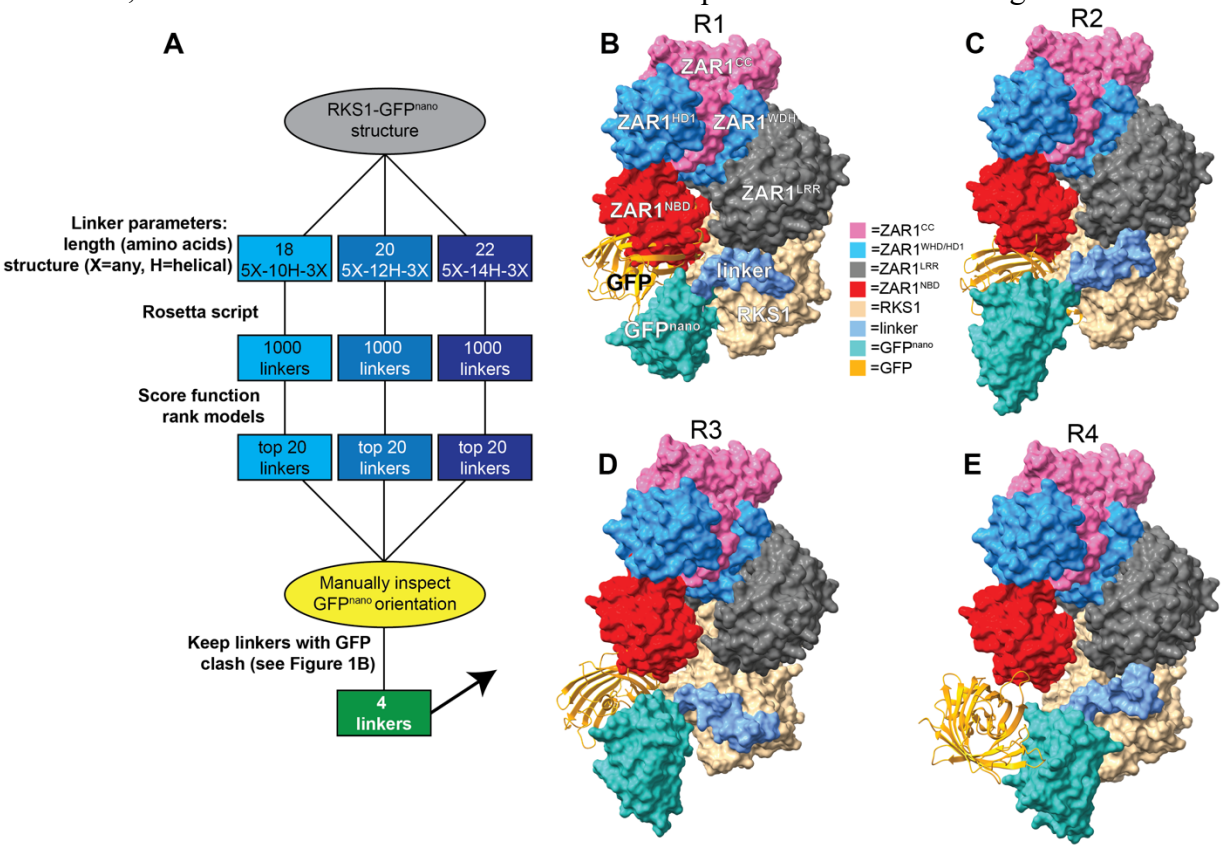



**Figure 3.** Approach for linker design with Rosetta and associated structural models. (A) Flowchart depicting our approach for Rosetta linker design for linkers of 18, 20 and 22 amino acids. (B) Structural models of RKS1-GFP$^{nano}$ fused by R1 (B), R2 (C), R3 (D) and R4 (E) linkers generated through our Rosetta pipeline (see Table 1) in complex with ZAR1 and GFP based on ZAR1-RKS1 structures and GFP$^{nano}$ GFP structures (Kubala et al., 2010; Wang et al., 2019a; Wang et al., 2019b). Images of structural models were produced using Chimera (Pettersen et al., 2004).

Rosetta did not allow for the generation of linkers of random length, or between a range of lengths. We therefore specified linkers of either 18, 20, or 22 amino acids. These lengths were determined based on the distance between our optimal position of the nanobody and the C-terminus of RKS1. To achieve a rigid linker, we specified most of the linker length as helical, with the remaining residues of random secondary structure. We generated 1000 linkers for each specified length and selected the top 20 scoring models, for a total of 60 models across 3 linker lengths. We visualized linkers using UCSF Chimera and identified four linkers out of our top 60 models that placed the GFP$^{nano}$ in approximately the correct orientation relative to ZAR1$^{NBD}$ (Figure 3,). We selected these four linkers for testing in planta (Table 1).

We transiently co-expressed each of our four RKS1- GFP$^{nano}$ constructs fused with Rosetta linkers with AtZAR1 and GFP in NbZAR1-VIGS *N. benthamiana*. We did not observe an HR from our RKS1-linker-nanobody constructs when we co-expressed AtZAR1 and GFP in *NbZAR1*-VIGS plants (Figure 2C, D). Silencing was effective because co-expression of ZED1 and HopZ1a in *NbZAR*- silenced plants did not trigger an HR. In addition, co-expression of AtZAR1, ZED1 and HopZ1a resulted in an HR as expected (Baudin et al., 2017).

*Modeling linker constructs with Alphafold*

Throughout this project, we sought to use new technological advances to improve the design of our linkers. The advent of Alphafold drastically improved the accuracy of protein prediction (Evans et al., 2022). We therefore modeled our random and Rosetta linker constructs using Alphafold. While Rosetta predicted that the four designed linkers placed GFP$^{nano}$ in the correct orientation relative to ZAR1$^{NBD}$, Alphafold predicted that all Rosetta linkers placed GFP$^{nano}$ oriented with the GFP binding interface away from ZAR1$^{NBD}$ and instead towards the RKS1 activation loop (Figure S1). It is possible that this orientation could also result in the activation of ZAR1, by displacing the RKS1 activation loop towards the ZAR1$^{NDB}$.

*Designing linkers with RFdiffusion, MPNN, and Alphafold*

Alphafold and other machine learning models have drastically advanced protein structure prediction. While diffusion models have enabled image and language generative AI, they have not been widely applied to protein design. RFdiffusion is the first deep-learning framework for general protein design, using a diffusion model to generate protein backbones with high accuracy (Watson et al., 2023). We sought to utilize RFdiffusion to generate RKS1-GFP$^{nano}$ linker backbones. Using LZerD, we docked RKS1 and GFP$^{nano}$ to determine energetically favorable conformations that placed proteins in the correct conformation relative to ZAR1$^{NBD}$ (Christoffer et al., 2021). We examined the conformation of GFP$^{nano}$ relative to RKS1 for the top 20 LZerD heterodimer models and identified one model that had RKS1 docked with GFP$^{nano}$ in the correct conformation relative to ZAR1$^{NBD}$ (Figure S2). This model, along with the R1 model from our Rosetta pipeline (Figure 3B) were utilized as templates for RFdiffusion linker backbone design (Figure 4A). The amino acids of RFdiffusion backbones were determined using ProteinMPNN and linkers were modeled with Alphafold2 (Dauparas et al., 2022; Evans et al., 2022; Wang et al., 2022). We selected the top 4 linkers based on ProteinMPNN scores and cloned them into the RKS1-GFP$^{nano}$ vector for transient expression in *N. benthamiana* (Figure 4, Table 1). As with all previously tested linkers, we did not observe an HR when the RFdiffusion linker constructs were co-expressed with AtZAR1 and GFP in *NbZAR1*-VIGS plants (Figure 2E, F). As expected, co-expression of AtZAR1, ZED1 and HopZ1a resulted in an HR, and HopZ1a and ZED1 did not trigger an HR in *NbZAR1*-silenced plants (Figure 2).

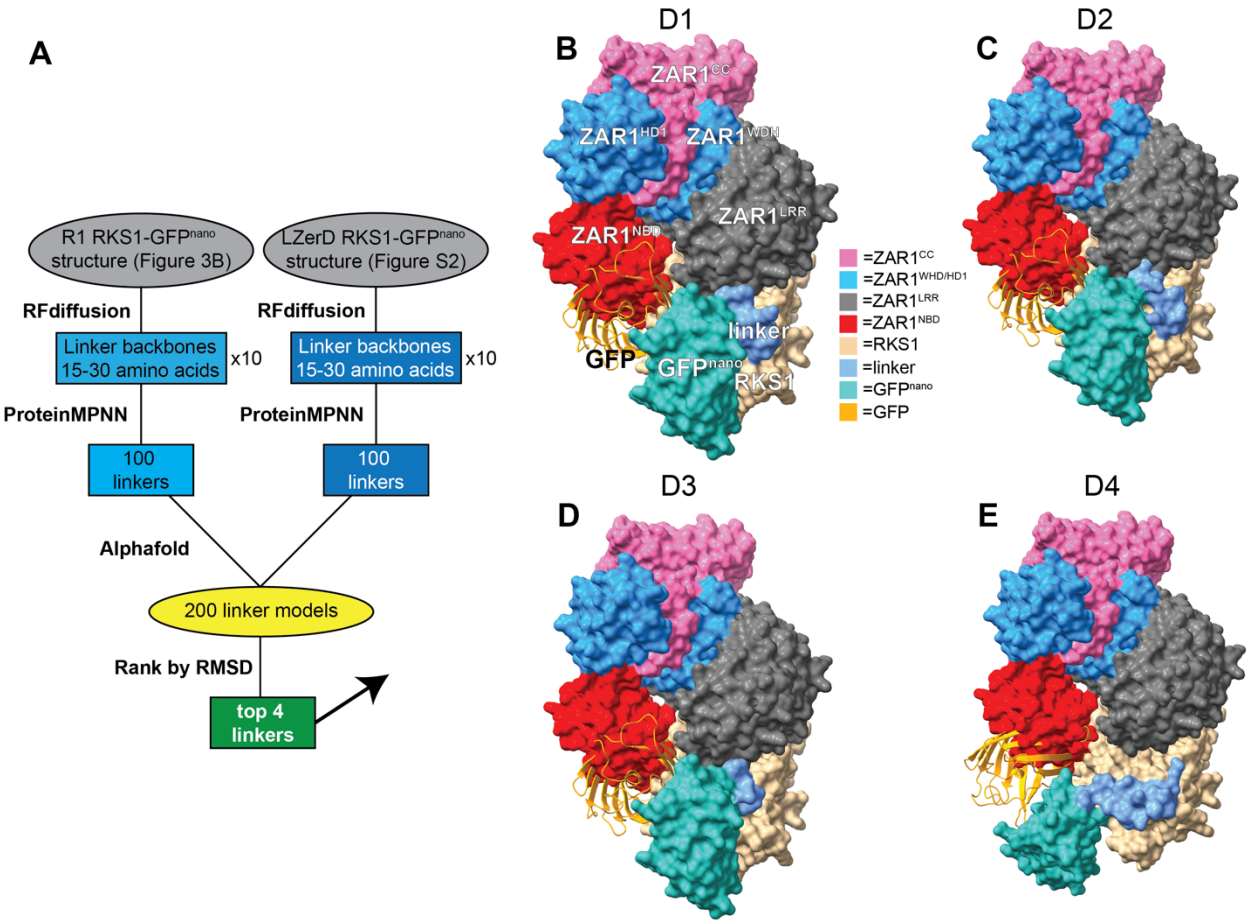



**Figure 4.** Approach for linker design using RFdiffusion and associated structural models. (A) Flowchart depicting our approach for RFdiffusion linker. (B) Structural models of RKS1-GFP$^{nano}$ fused by D1 (B), D2 (C), D3 (D) and D4 (E) linkers generated through our RFdiffusion pipeline (see Table 1) in complex with ZAR1 and GFP based on ZAR1-RKS1 structures and GFP$^{nano}$ GFP structures (Kubala et al., 2010; Wang et al., 2019a; Wang et al., 2019b). Images of structural models were produced using Chimera (Pettersen et al., 2004).

*RKS1-GFP$^{nano}$ proteins are expressed, and interact with ZAR1 and GFP*

To determine if poor RKS1-GFP$^{nano}$ protein accumulation caused an absence of the HR, we conducted western blot assays to check for RKS1-linker-GFP$^{nano}$ protein. We transiently expressed each linker construct in *NbZAR1*-VIGS *N. benthamiana* and collected tissue samples 24 hours after dexamethasone induction. We detected protein accumulation at levels similar to RKS1$^{WT}$ for all constructs except L1, L2, R3, and D2 (Figure 5A).

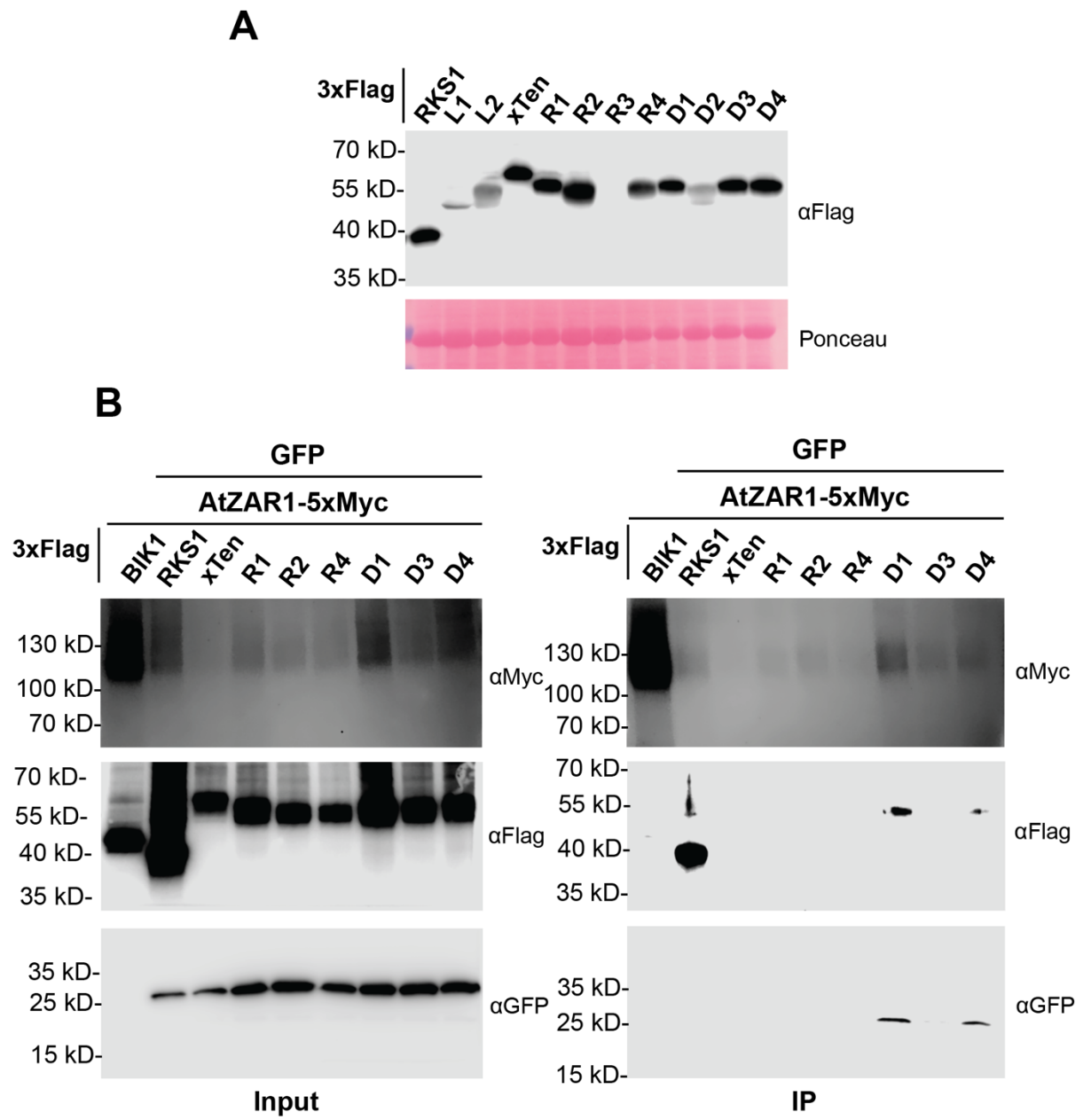



**Figure 5**. RKS1-GFP$^{nano}$ fusions are expressed and form a complex with ZAR1 and GFP. (A) Anti-Flag western blot showing that most RKS1-GFP$^{nano}$ fusions are expressed in *N. benthamiana* (see Table 1). RKS1-GFP$^{nano}$ fusions were transiently expressed in *N. benthamiana NbZAR1*-VIGS and tissue samples were collected 24 hours post-dexamethasone induction. Ponceau stain is included to show equal loading. (B) Representative western blots of AtZAR1 coimmunopurification with RKS1-GFP$^{nano}$ in *N. benthamiana*. Each RKS1-GFP$^{nano}$ fusion was cloned as an in-frame fusion with a C-terminal 3xFlag tag in a dexamethasone-inducible expression vector and co-expressed with AtZAR1 cloned as an in-frame fusion with a C-terminal 5x-Myc tag. AtZAR1 was immunoprecipitated using anti-Myc beads. Coimmunoprecipitation experiments were repeated twice with similar results. Predicted molecular masses of the proteins are AtZAR1-5xMyc 107 kD; BIK1-3xFlag 47kD, RKS1-3xFlag 43 kD, RKS1-GFP$^{nano}$-3xFlag 55-57 kD, GFP 27 kD.

To determine if our RKS1-linker-GFP$^{nano}$ proteins interacted with ZAR1 and GFP, we conducted co-immunoprecipitation (co-IP) assays. We transiently expressed each RKS1-linker-GFP$^{nano}$ construct as an in-frame fusion to a 3xFlag tag, ZAR1 as an in-frame fusion to a 5xMyc

tag, and GFP in *NbZAR1*-VIGS *N. benthamiana*. We omitted L1, L2, R3, and D2 as these were weakly expressed (Figure 5A). We collected tissue samples 24 hours after induction and immunoprecipitated samples with magnetic Myc beads. We noted a consistent strong interaction between ZAR1 and RKS1, but not between ZAR1 and BIK1, as previously demonstrated (Figure 5B) (Diplock et al., 2023). We observed a consistent weak interaction between ZAR1 and D1, or ZAR1 and D4, relative to ZAR1 and RKS1 (Figure 5B). Moreover, this weak interaction was also associated with co-immunoprecipitation of GFP, suggesting that ZAR1-D1-GFP and ZAR1-D4-GFP form a complex. We did not observe consistent AtZAR1-GFP complex formation resulting from the xTen, R1, R2, R4, or D3 linkers (Figure 5B). To predict if 3xFlag fusions present at the C-terminus of all RKS1-GFP$^{nano}$ fusions could be inhibiting ZAR1 or GFP interaction, we modeled the GFP$^{nano}$-3xFlag fusion with Alphafold and overlayed or R1 model. The 3xFlag was not predicted to clash with ZAR1 or GFP (Figure S3).

**Discussion**

Here, we attempted to selectively activate ZAR1 in the presence of an antigen through RKS1-GFP$^{nano}$ fusions (Figure 1). Using a variety of linker design methods (Table 1, Figure 3,4), we were unable to successfully achieve activation in the presence of GFP with any of the RKS1-GFP$^{nano}$ fusions (Figure 2). We confirmed that most of our tested RKS1-GFP$^{nano}$ fusions were expressed, and that some facilitated interaction between ZAR1 and GFP (Figure 5). While we were unable to achieve immune activation of ZAR1, our linker design methods may be of use for other applications, and activation of ZAR1 through RKS1-GFP$^{nano}$ fusions may be possible as new protein design approaches become available. Achieving activation of ZAR1 through a guardee-nanobody fusion could enable rapid generation of plant immunity to custom pathogen targets using native ZAR1 orthologs across angiosperms.

We have several hypotheses for why ZAR1 activation was not achieved. Broadly, it is possible that no tested linkers achieved the correct orientations between RKS1, GFP$^{nano}$, GFP, and ZAR1$^{NBD}$. Alternatively, it is possible that linkers achieved correct protein domain orientations but could not activate ZAR1. Previous research has shown that ZRKs, including RKS1 interact with ZAR1 in the absence of effectors (Wang et al., 2015; Diplock et al., 2023). While we hypothesized that GFP binding to RKS1-GFP$^{nano}$ would result in ZAR1 activation, it is possible that GFP could not access the binding interface, due to its globular nature. Therefore, it is possible that, while GFP$^{nano}$ was in the correct orientation relative to ZAR1$^{NBD}$, the RKS1-ZAR1 interaction prevented GFP from accessing the nanobody binding interface. While we observed ZAR1, RKS1-GFP$^{nano}$, GFP complex formation for D1 and D4 (Figure 5B), it is possible that these linkers enabled complex formation because they placed GFP$^{nano}$ in an orientation away from ZAR1$^{NBD}$. Alternatively, it is possible that protein design approaches were not accurate enough to generate linkers with the correct GFP$^{nano}$ orientation. A possible solution to overcome this barrier would be to test a greater number of linkers. We have previously shown that weakening ZAR1-ZRK interactions results in a loss of the HR (Baudin et al., 2017; Diplock et al., 2023). We observed weaker interactions between ZAR1 and RKS1-GFP$^{nano}$ when compared to ZAR and RKS1 (Figure 5B), suggesting perturbed ZAR1 interaction could explain the absence of HR activation. Finally, it is possible that epitope tags inhibited activation of ZAR1 or correct orientation of GFP$^{nano}$ relative to ZAR1$^{NBD}$. All RKS1-GFP$^{nano}$ fusions were cloned with a 3xFlag epitope tag, which was not modeled in any of our design programs. It is possible that this epitope tag inhibited the accuracy of our protein modeling. However, we feel it is unlikely that a 3xFlag epitope tag would influence GFP-GFP$^{nano}$ interaction, given the C-

terminus of the GFP$^{nano}$ is opposite the GFP binding interface (Kubala et al., 2010). Furthermore, when we modeled the GFP$^{nano}$-3xFlag fusion with Alphafold and overlayed or R1 model, the 3xFlag was not predicted to clash with ZAR1 or GFP (Figure S3).

Although autoactivation of some ZAR1 orthologs has been demonstrated, this has not been achieved with AtZAR1 in *N. benthamiana*. Transient overexpression of certain NLRs has been shown to result in HR in *N. benthamiana* in the absence of pathogen effectors (Day et al., 2005). However, this does not occur with AtZAR1 or NbZAR1 (Lewis et al., 2013; Baudin et al., 2017). Several mutations in key NLR regulatory motifs have been shown to result in autoactivation, especially substitution of a valine for aspartate in the MHD motif (Bendahmane et al., 2002; Howles et al., 2005; De La Fuente Van Bentem et al., 2005; Williams et al., 2011; Roberts et al., 2013). While a NbZAR1 MHD mutant was autoactive upon transient expression in *N. benthamiana* (Harant et al., 2022), we have previously shown that the corresponding mutation in AtZAR1 does not result in autoactivity in *N. benthamiana* (Baudin et al., 2019). Moreover, while transient expression of certain ZRKs including RKS1 results in autoactivation of NbZAR1 in the absence of pathogen effectors, no autoactivation of AtZAR1 has been observed when it is co-expressed with ZRKs (Gong et al., 2022; Diplock et al., 2023). Overall, these findings point towards NbZAR1 activation being less regulated and/or easier to achieve than AtZAR1. While we do not have a molecular explanation for this trend, it is a potential explanation why co-expressing AtZAR1 with RKS1-GFP$^{nano}$ fusions and GFP did not result in the HR. We selected RKS1 as our ZAR1 guardee for fusions due to the availability of structural data and overcame the autoactive phenotype between NbZAR1 and RKS1 by using *NbZAR1*-VIGS plants. However, it is possible that utilizing a ZRK that is not autoactive with NbZAR1 could achieve NbZAR1 activation when fused to GFP$^{nano}$ and co-expressed with GFP. Finally, AtZAR1 has been shown to mediate immunity with and without the presence of the HR (Breit-McNally et al., 2022). While the mechanism behind this ability is not well understood, it is possible that RKS1-GFP$^{nano}$ fusions activated ZAR1-mediated resistance but not ZAR1-mediated HR. This could be investigated through RKS1-GFP$^{nano}$ stable transgenics or transient expression in *N. benthamiana* followed by *P. syringae* infection (Buscaill et al., 2020).

Although we were unable to achieve antibody-mediated activation of plant immunity through ZAR1, our novel approach of incorporating nanobodies into NLR guardees may inspire other novel approaches for achieving recognition of emerging pathogens in plants. One NLR independent approach could be the utilization of antibody-mediated degradation systems to target effectors of virulent pathogens. In such a system, antibodies could be raised against pathogen virulence targets to target them for ubiquitination and degradation (Cotton et al., 2021), which could result in decreased pathogen virulence. Alternatively, nanobodies could be used to outcompete binding of pathogen proteins to host targets, as has been in mammalian systems (Huo et al., 2021). Other systems exploring nanobody fusions directly to NLRs, similar to the NLR-nanobody fusions that enabled activation of the Pik-1 in the presence of the nanobody antigen in rice (Kourelis et al., 2023), could further expand plant effector recognition repertoires. Moreover, our linker design methodology contributes to protein design efforts outside of plant systems. As the field of rational protein design continues to expand, new applications will emerge for custom protein linker designs.

**References**

Adachi H, Sakai T, Kourelis J, Pai H, Gonzalez Hernandez JL, Utsumi Y, Seki M, Maqbool A, Kamoun S (2023). Jurassic NLR: Conserved and dynamic evolutionary features of the

atypically ancient immune receptor ZAR1. *Plant Cell*, (*1998*), 1–50
Amet N, Lee HF, Shen WC (2009). Insertion of the designed helical linker led to increased expression of Tf-based fusion proteins. *Pharm Res*, *26*(*3*), 523–528
Baudin M, Hassan JA, Schreiber KJ, Lewis JD (2017). Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol*, *174*(*4*), 2038–2053
Baudin M, Schreiber KJ, Martin EC, Petrescu AJ, Lewis JD (2019). Structure–function analysis of ZAR1 immune receptor reveals key molecular interactions for activity. *Plant J*, *101*(*2*), 352–370
Bendahmane A, Farnham G, Moffett P, Baulcombe DC (2002). Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the *Rx*locus of potato. *Plant J*, *32*(*2*), 195–204
Bi G, Su M, Li N, Liang Y, Dang S, Xu J, Hu M, Wang J, Zou M, Deng Y, et al (2021). The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. *Cell*, 1–14
Breit-McNally C, Laflamme B, Singh RA, Desveaux D, Guttman DS (2022). ZAR1: Guardian of plant kinases. *Front Plant Sci*, *13*(*September*), 1–8
Buscaill P, Sanguankiattichai N, Lee YJ, Kourelis J, Preston G, van der Hoorn RAL (2020). Agromonas: a rapid disease assay for Pseudomonas syringae growth in agroinfiltrated leaves. *bioRxiv*, 2020.08.10.243808
Chen X, Zaro JL, Shen W-C (2013). Fusion protein linkers: Property, design and functionality. *Adv Drug Deliv Rev*, *65*(*10*), 1357–1369
Christoffer C, Chen S, Bharadwaj V, Aderinwale T, Kumar V, Hormati M, Kihara D (2021). LZerD webserver for pairwise and multiple protein-protein docking. *Nucleic Acids Res*, *49*(*W1*), W359–W365
Cotton AD, Nguyen DP, Gramespacher JA, Seiple IB, Wells JA (2021). Development of Antibody-Based PROTACs for the Degradation of the Cell-Surface Immune Checkpoint Protein PD-L1. *J Am Chem Soc*, *143*(*2*), 593–598
Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, Wicky BIM, Courbet A, de Haas RJ, Bethel N, et al (2022). Robust deep learning–based protein sequence design using ProteinMPNN. *Science*, *378*(*6615*), 49–56
Day B, Dahlbeck D, Huang J, Chisholm ST, Li D, Staskawicz BJ (2005). Molecular basis for the RIN4 negative regulation of RPS2 disease resistance. *Plant Cell*, *17*(*4*), 1292–1305
Diplock N, Baudin M, Harden L, Silva CJ, Erickson-Beltran ML, Hassan JA, Lewis JD (2023). Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1. *Plant Cell Environ*, *46*(*7*), 2238–2254
Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, et al (2022). Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.10.04.463034
George RA, Heringa J (2002). An analysis of protein domain linkers: Their classification and role in protein folding. *Protein Eng*, *15*(*11*), 871–879
Gokhale RS, Khosla C (2000). Role of linkers in communication between protein modules. *Curr Opin Chem Biol*, *4*(*1*), 22–27
Gong Z, Qi J, Hu M, Bi G, Zhou J-M, Han G-Z (2022). The origin and evolution of a plant resistosome. *Plant Cell*, *34*(*5*), 1600–1620
Harant A, Pai H, Sakai T, Kamoun S, Adachi H (2022). A vector system for fast-forward studies

of the HOPZ-ACTIVATED RESISTANCE1 (ZAR1) resistosome in the model plant Nicotiana benthamiana. *Plant Physiol*, *188*(*1*), 70–80
Heath MC (2000). Hypersensitive response-related death. *Plant Mol Biol*, *44*(*3*), 321–334
Howles P, Lawrence G, Finnegan J, McFadden H, Ayliffe M, Dodds P, Ellis J (2005). Autoactive alleles of the flax L6 rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. *Mol Plant-Microbe Interact*, *18*(*6*), 570–582
Huo J, Bas A Le, Ruza RR, Duyvesteyn HME, Mikolajek H, Malinauskas T, Tan TK, Rijal P, Dumoux M, Ward PN, et al (2021). Author Correction: Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2 (Nature Structural & Molecular Biology, (2020), 27, 9, (846-854), 10.1038/s41594-020-0469-6). *Nat Struct Mol Biol*, *28*(*3*), 326
Ikebe M, Kambara T, Stafford WF, Sata M, Katayama E, Ikebe R (1998). A hinge at the central helix of the regulatory light chain of myosin is critical for phosphorylation-dependent regulation of smooth muscle myosin motor activity. *J Biol Chem*, *273*(*28*), 17702–17707
Jones JDG, Dangl JL (2006). The plant immune system. *Nature*, *444*(*7117*), 323–329
Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, *10*(*6*), 845–858
Kourelis J, Marchal C, Posbeyikian A, Harant A, Kamoun S (2023). NLR immune receptor–nanobody fusions confer plant disease resistance. *Science*, *379*(*6635*), 934–939
Kubala MH, Kovtun O, Alexandrov K, Collins BM (2010). Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. *Protein Sci*, *19*(*12*), 2389–2401
Kuhlman B, Jacobs T, Linskey T (2016). Computational design of protein linkers. *Methods Mol Biol*, *1414*341–351
De La Fuente Van Bentem S, Vossen JH, De Vries KJ, Van Wees S, Tameling WIL, Dekker HL, De Koster CG, Haring MA, Takken FLW, Cornelissen BJC (2005). Heat shock protein 90 and its co-chaperone protein phosphatase 5 interact with distinct regions of the tomato I-2 disease resistance protein. *Plant J*, *43*(*2*), 284–298
Lewis JD, Lee AHY, Hassan JA, Wana J, Hurleya B, Jhingree JR, Wang PW, Lo T, Youn JY, Guttman DS, et al (2013). The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the Pseudomonas syringae type III effector HopZ1a. *Proc Natl Acad Sci USA*, *110*(*46*), 18722–18727
Lewis JD, Wu R, Guttman DS, Desveaux D (2010). Allele-specific virulence attenuation of the Pseudomonas syringae HopZ1a type III effector via the Arabidopsis ZAR1 resistance protein. *PLoS Genet*, *6*(*4*), e1000894
Maeda Y, Ueda H, Kazami J, Kawano G, Suzuki E, Nagamune T (1997). Engineering of functional chimeric protein G-Vargula luciferase. *Anal Biochem*, *249*(*2*), 147–152
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004). UCSF Chimera-A visualization system for exploratory research and analysis. *J Comput Chem*, *25*(*13*), 1605–1612
Roberts M, Tang S, Stallmann A, Dangl JL, Bonardi V (2013). Genetic requirements for signaling from an autoactive plant NB-LRR intracellular innate immune receptor. *PLoS Genet*, *9*(*4*), e1003465
Wang G, Roux B, Feng F, Guy E, Li L, Li N, Zhang X, Lautier M, Jardinaud MF, Chabannes M, et al (2015). The decoy substrate of a pathogen effector and a pseudokinase specify pathogen-induced modified-self recognition and immunity in plants. *Cell Host Microbe*,

*18*(*3*), 285–295
Wang J, Hu M, Wang J, Qi J, Han Z, Wang G, Qi Y, Wang HW, Zhou JM, Chai J (2019a). Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science*, *364*(*6435*), eaav5870
Wang J, Lisanza S, Juergens D, Tischer D, Watson JL, Castro KM, Ragotte R, Saragovi A, Milles LF, Baek M, et al (2022). Scaffolding protein functional sites using deep learning. *Science*, *377*(*6604*), 387–394
Wang J, Wang J, Hu M, Wu S, Qi J, Wang G, Han Z, Qi Y, Gao N, Wang HW, et al (2019b). Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science*, *364*(*6435*), eaav5868
Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, Ahern W, Borst AJ, Ragotte RJ, Milles LF, et al (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, *620*(*7976*), 1089–1100
Williams SJ, Sornaraj P, DeCourcy-Ireland E, Menz RI, Kobe B, Ellis JG, Dodds PN, Anderson PA (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Mol Plant-Microbe Interact*, *24*(*8*), 897–906

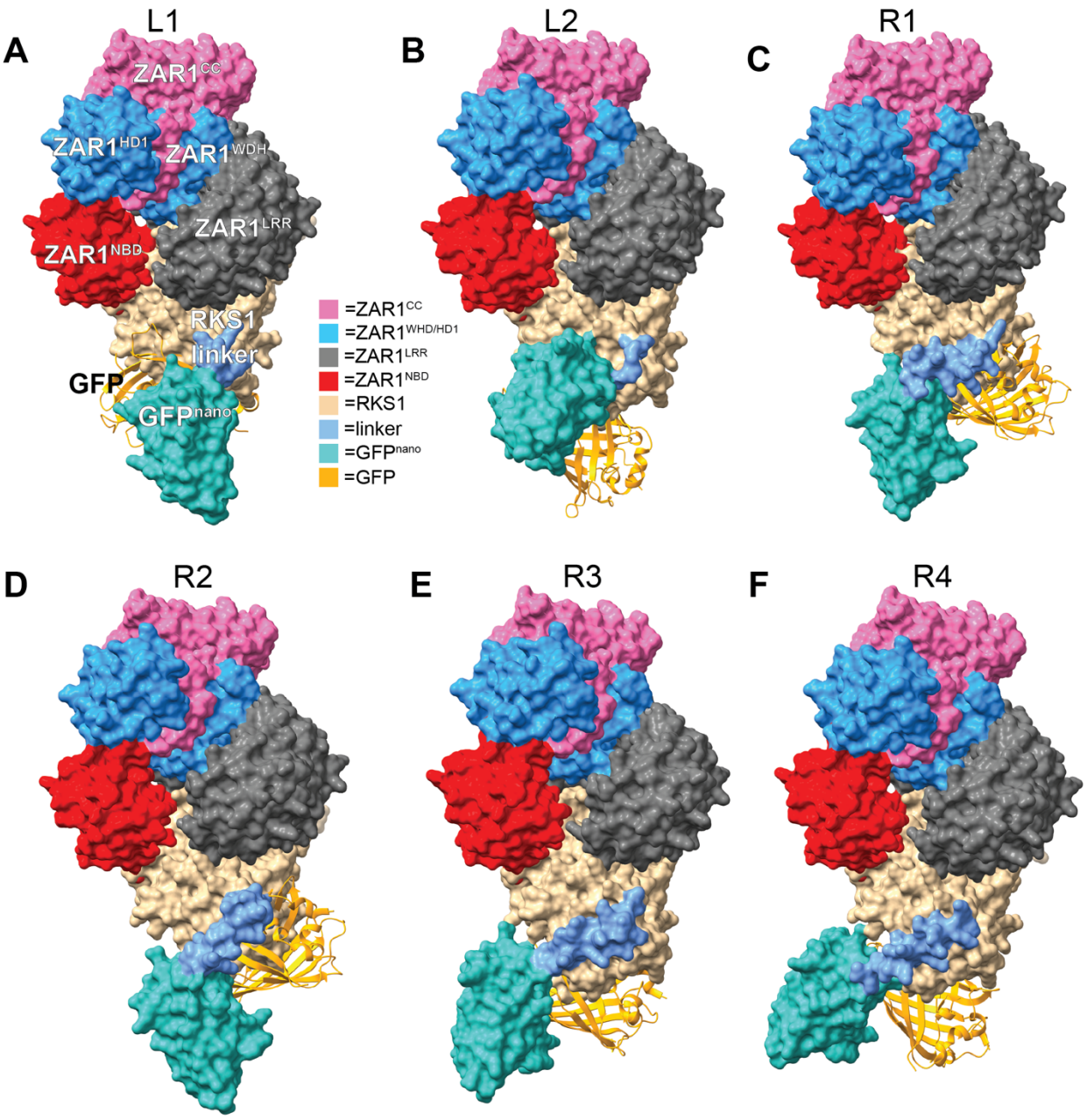



**Figure S1**. Alphafold models of RKS1-GFP$^{nano}$ fusions with Random and Rosetta linkers. Top Alphafold structural models of RKS1-GFP$^{nano}$ fused by (A) L1, (B) L2, (C) R1, (D) R2, (E) R3, and (D) R4 (see Table 1) in complex with ZAR1 and GFP based on ZAR1-RKS1 structures and GFP$^{nano}$ GFP structures (Kubala et al., 2010; Wang et al., 2019a; Wang et al., 2019b). Images of structural models were produced using Chimera (Pettersen et al., 2004).

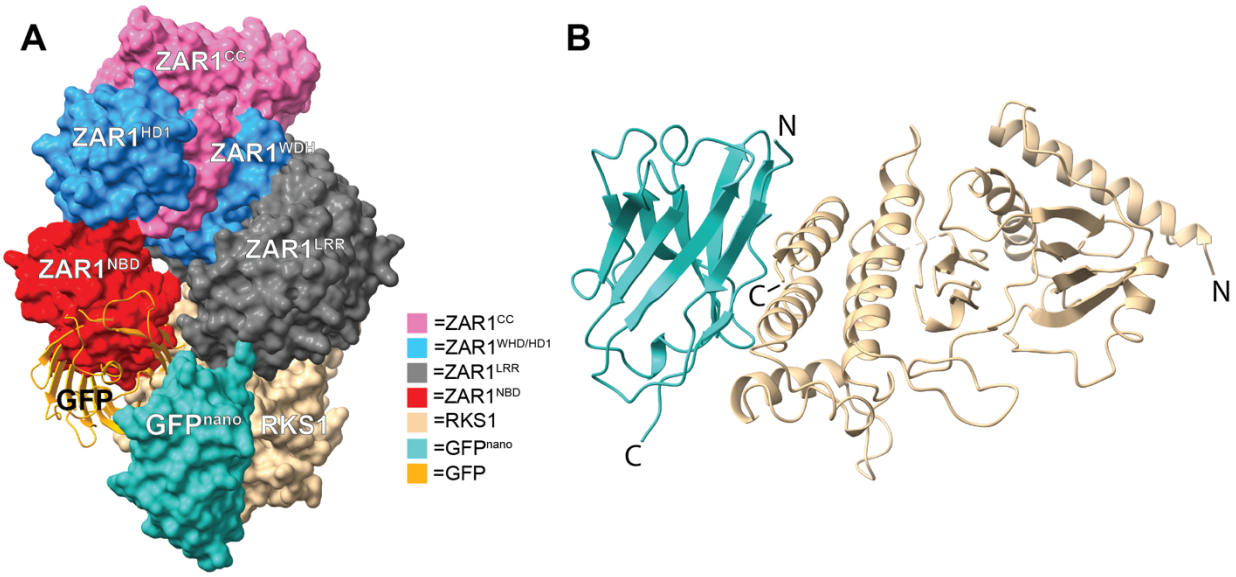



**Figure S2.** Protein-protein docking of RKS1 and GFP$^{nano}$ using LZerD. (A) Top LZerD model of RKS1 docked with GFP$^{nano}$ in complex with ZAR1 and GFP based on ZAR1-RKS1 structures and GFP$^{nano}$ GFP structures (Kubala et al., 2010; Wang et al., 2019a; Wang et al., 2019b). (B) LZerD of RKS1 (beige) docked with GFP$^{nano}$ (blue). Images of structural models were produced using Chimera (Pettersen et al., 2004).

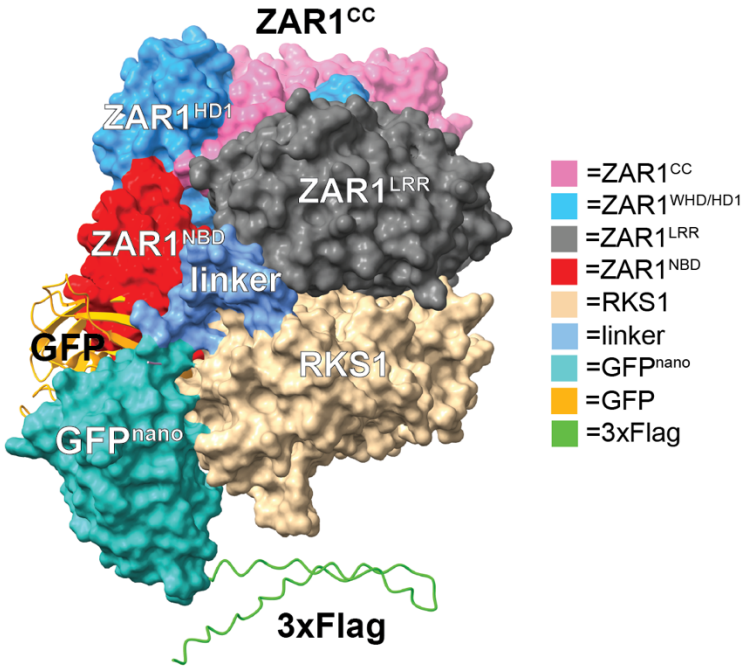



**Figure S3.** Rosetta R1 model with Alphafold prediction of 3xFlag. The GFP$^{nano}$-3xFlag fusion was modeled in Alphafold and overlayed with the Rosetta R1 model in Chimera (see Figure 3 for R1 model). Images of structural models were produced using Chimera (Pettersen et al., 2004).

## CONCLUSION

While the approach presented here was unsuccessful, incorporating adaptive immunity components into plant innate immunity recognition pathways remains a promising approach for rational design of pathogen recognition, including recognition of *P. syringae*.

**APPENDIX I:** *A. thaliana* Wall Associated Kinase-Like proteins are targeted by HopZ1a

## PREFACE

The work presented in this appendix was a collaboration with Derek Xiang and Jennifer D. Lewis. I performed the expression analysis. I performed the yeast-two hybrids, BiFC assays, and transient expression HR assays with Derek Xiang. I wrote the appendix with assistance from Jennifer D. Lewis. The work has been reproduced below with the permission of all authors.

Investigating the natural diversity of Wall Associated Kinase-Like proteins targeted by the *P. syringae* effector HopZ1a is a key step towards rational design of effector recognition through ZAR1.

### Background

ZED1 and ZED1-related kinases likely evolved from the wall-associated kinase (WAK) and wall-associated kinase-like (WAKLs) proteins through the loss of WAKL extracellular domains[1]. Some WAKs have previously described roles in PTI in diverse plant species[2,3]. We therefore hypothesized that HopZ1a may target WAKLs to suppress PTI, while activating resistance through the closely related ZRK, ZED1. *A. thaliana* has 22 WAKLs and 5 WAKs[4]. All WAKs and WAKL1-7 are present in gene clusters (Figure 1, Table 1), similar to the arrangement of *A. thaliana* ZRKs[4,5].

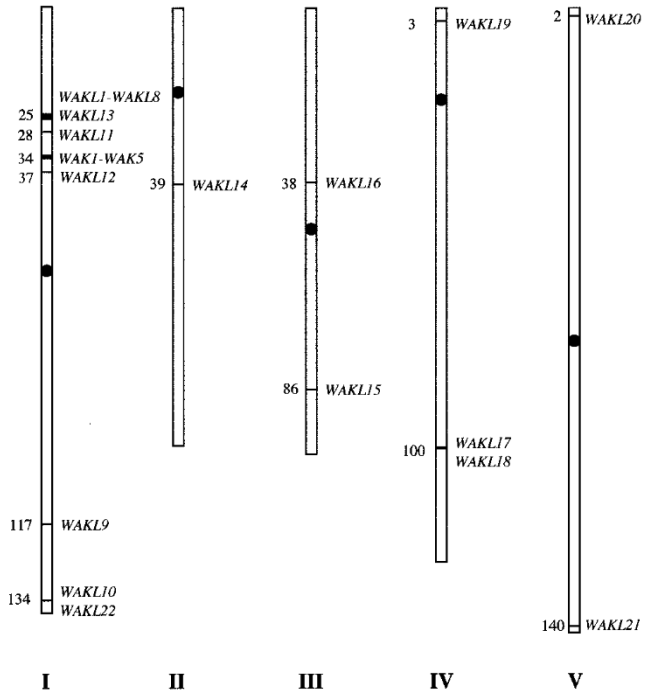



**Figure 1.** Genomic map of *A. thaliana* WAKs and WAKLs from Verica and He (2002)[4]. Vertical bars indicate *A. thaliana* chromosomes, horizontal bars indicate locations of genes, and numbers indicate position of genes on each chromosome[4].

**Table 1**. WAK and WAKL genes in *A. thaliana* from Verica and He (2002)[4]

| Gene | Location |
|---|---|

| | |
|---|---|
| WAK1 | At1g21250 |
| WAK2 | At1g21270 |
| WAK3 | At1g21240 |
| WAK4 | At1g21210 |
| WAK5 | At1g21230 |
| WAKL1 | At1g16120 |
| WAKL2 | At1g16130 |
| WAKL3 | At1g16140 |
| WAKL4 | At1g16150 |
| WAKL5 | At1g16160 |
| WAKL6 | At1g16110 |
| WAKL7 | At1g16090 |
| WAKL8 | At1g16260 |
| WAKL9 | At1g69730 |
| WAKL10 | At1g79680 |
| WAKL11 | At1g19390 |
| WAKL12 | At1g22720 |
| WAKL13 | At1g17910 |
| WAKL14 | At2g23450 |
| WAKL15 | At3g53840 |
| WAKL16 | At3g25490 |
| WAKL17 | At4g31100 |
| WAKL18 | At4g31110 |
| WAKL20 | At5g02070 |
| WAKL21 | At5g66790 |
| WAKL22 | At1g79670 |

**Methods**

The kinase domains of WAKL1, 2, 5, and 6, were amplified from genomic DNA while the kinase domains for WAKL4, 8, 10 and 22 were amplified from plasmids from the ABRC. WAKL amino acid sequences were aligned, and kinase domains were defined relative WAKL2$^{KIN}$ which has previously been shown to interact with ZAR1[6]. WAKLs were amplified to contain either attB sites or 5'*Xho*I and 3'*Eco*RI sites. Fragments were cloned into pDONR207 using the BP reaction, or pJG4-5 via restriction enzyme digest. WAKL pDONR207 clones were transferred into a previously constructed Gateway compatible pDex-cYFP vector via the LR reaction (see Chapter 4). Methods for plant growth conditions, transient expression, and HR assays were the same as described in Chapter 3. Methods for yeast interaction assays and BiFC assays were the same as described in Chapter 4.

**Results**

*Specific WAKLs are induced by flg22*

A recent study described the role of tomato WAK1 in *P. syringae* PTI[3]. Expression of tomato WAK1 was upregulated in response to the *P. syringae* PAMP, flg22. To determine if *A. thaliana* WAKs or WAKLs were upregulated in response to flg22, we used expression data available through the BAR website. We excluded WAKL3 which is a pseudogene and WAKL7 which has a truncated kinase domain. Several WAKLs did not have expression data for flg22

treatment and were excluded. WAKL2, 4, 11, 14, 20, and 22 were upregulated following flg22 treatment while no WAKs had differential expression after flg22 treatment (Figure 2).

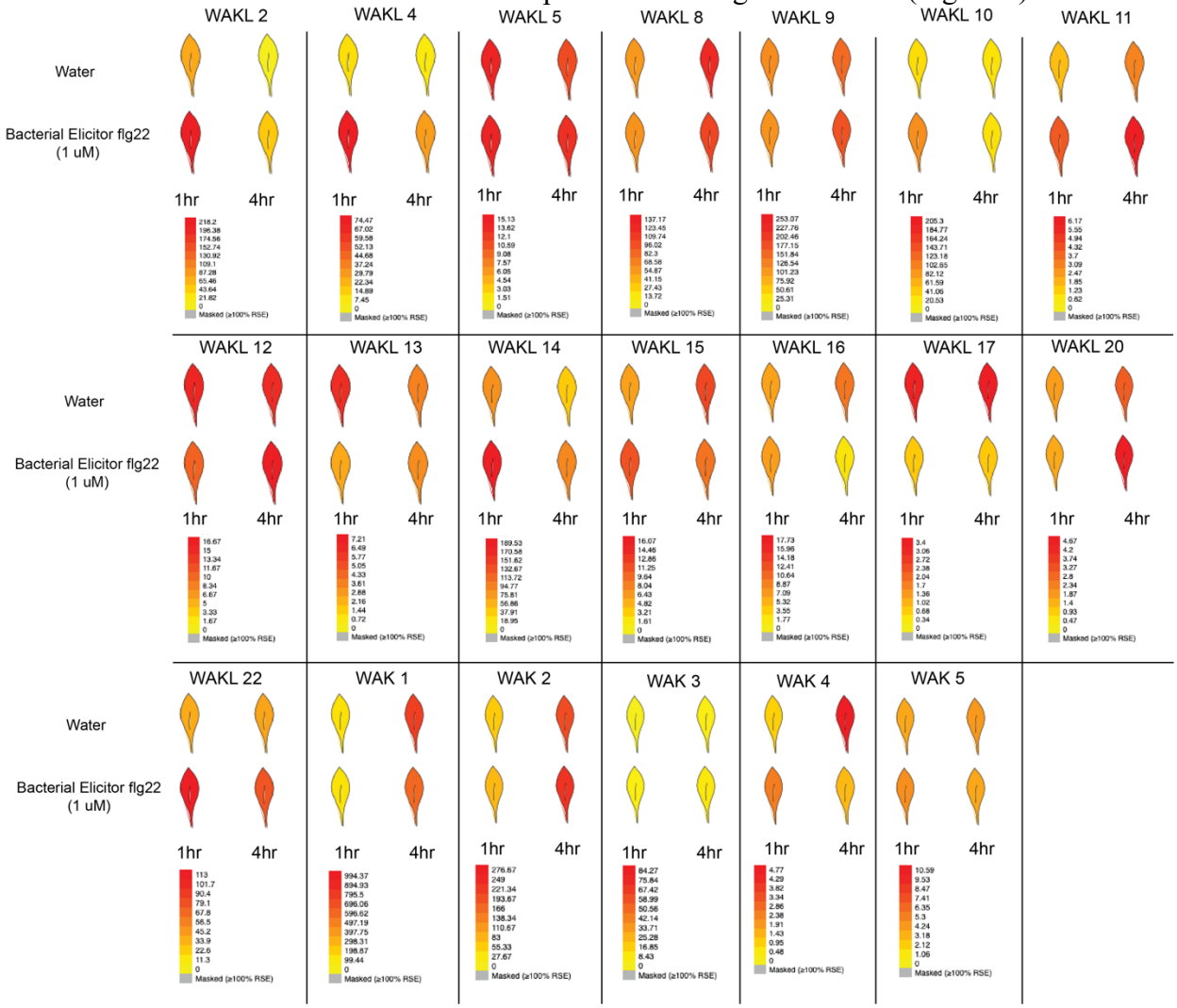



**Figure 2.** WAKLs are upregulated following flg22 treatment. Data was collected from a subset of biotic elicitors on the BAR website[7]. Images were modified from the BAR website[7].

*HopZ1a interacts with the kinase domain of some WAKLs in planta*

We hypothesized that the WAKLs most likely to be targeted by HopZ1a would be most related in sequence to ZED1. WAKL2 is the most similar kinase domain to ZED1 outside of the ZRK family determined by BLAST search[8]. WAKL2 is in a genomic cluster containing WAKL1-7, which are phylogenetically related to WAKL2 (Figure 1)[4]. WAKL22 was included since it has a previously described role in resistance to the fungal pathogens *Fusarium oxysporum*[9]. To test if HopZ1a targets the kinase domain of WAKLs, we cloned each WAKL kinase domain as an in-frame fusion with cYFP, and transiently co-expressed each with HopZ1a$^{C/A}$ cloned as an in-frame fusion to nYFP. We noted a strong interaction between HopZ1a$^{C/A}$ and the kinase domain of WAKL1, 2, 4, and 8 (Figure 3A). To better quantify WAKL-HopZ1a interactions, we quantified and compared fluorescence intensity (Figure 3B).

WAKL2, 4, and 8 interacted with HopZ1a at roughly the same strength as ZED1, while WAKL1 interacted with HopZ1a at greater strength (Figure 3B).

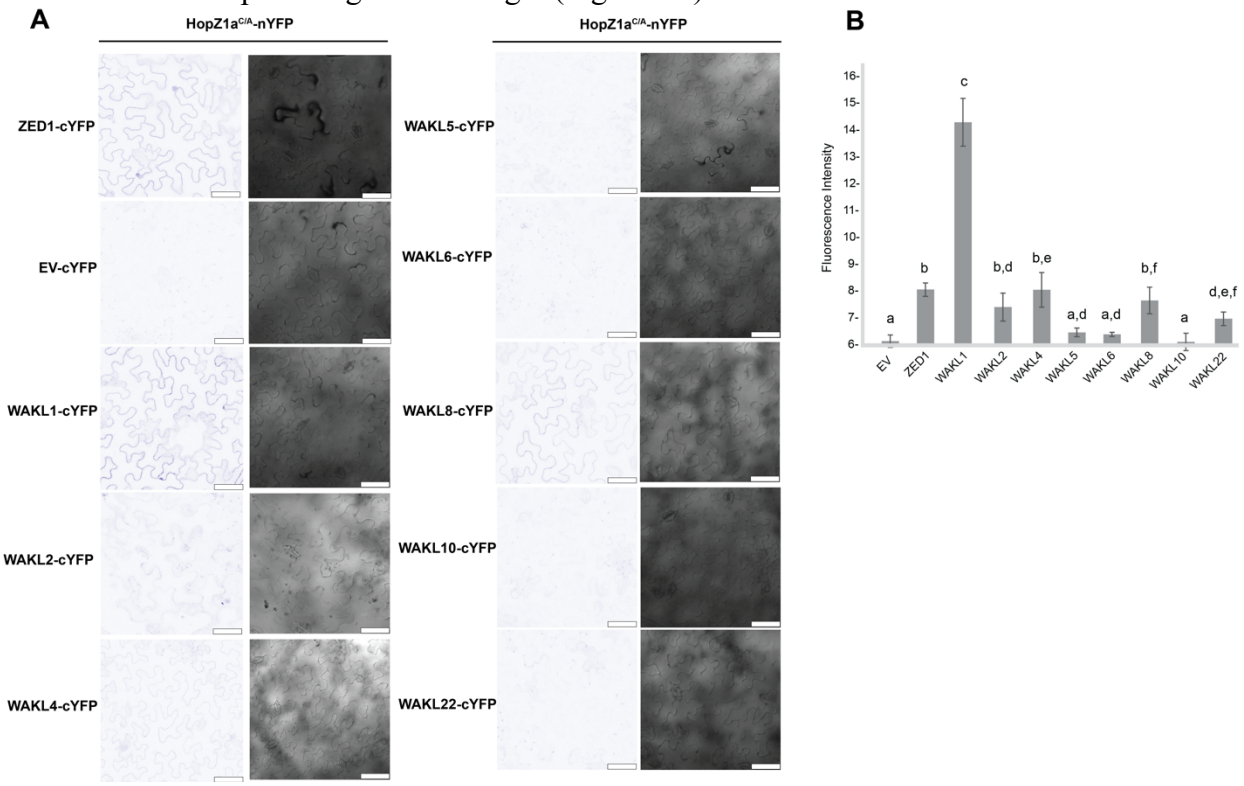



**Figure 3.** Kinase domains of WAKLs interact with HopZ1a. (A) Representative images showing the inverted YFP channel (top) and YFP and brightfield channel merged (bottom) from bimolecular fluorescence complementation (BiFC) assays showing interaction between specific WAKLs and HopZ1a$^{C216A}$. All WAKLs were cloned as in-frame fusions with cYFP and were co-expressed with HopZ1a$^{C216A}$ fused in-frame with nYFP in *N. benthamiana*. Images were captured 24 hours after dexamethasone induction using a Zeiss LSM710 confocal scanning microscope. White bar indicates 50μm. (B) Mean fluorescence intensity from all fields imaged in BiFC experiments was calculated using the Leica Application Suite X software (n=12-13 images across 6 plants and 3 experiments). Error bars represent standard error of means. T-tests were conducted to determine statistical significance, indicated by letters above bars ($p<0.05$).

*HopZ1a does not interact with WAKLs in yeast*

To determine if WAKLs could interact with HopZ1a in a heterologous system, we tested each WAKL for interaction with HopZ1a$^{C/A}$ in yeast. While ZED1 was able to interact with HopZ1a, no WAKLs could interact with HopZ1a in yeast (Figure 4). HopF2 and ShcF2 interacted strongly as expected[10].

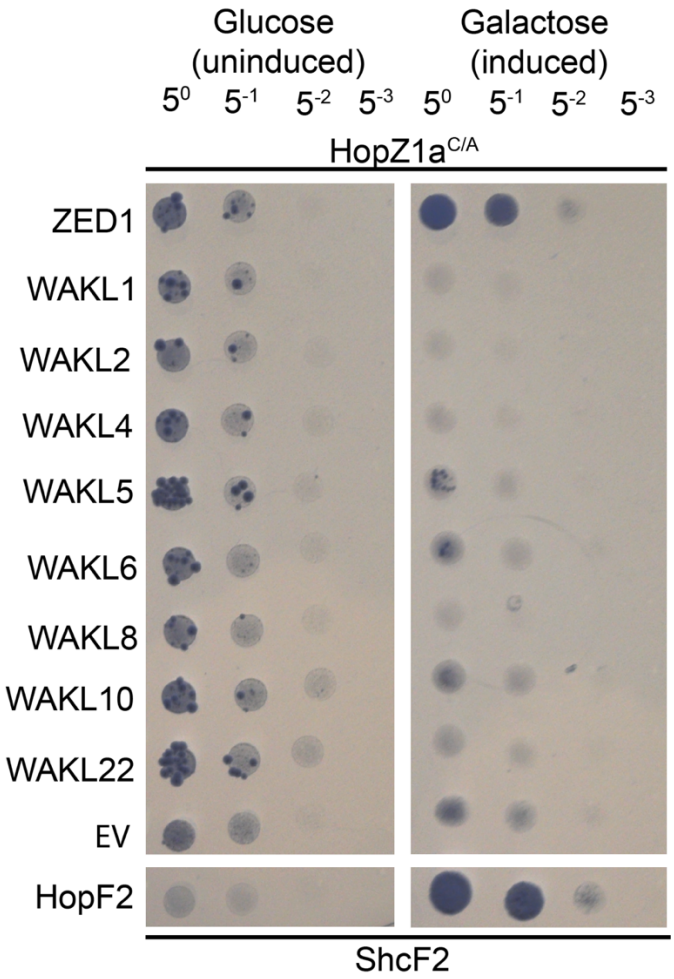



**Figure 4.** Kinase domains of WAKLs do not interact with HopZ1a in yeast. Kinase domains of WAKLs were cloned as an in-frame fusion to the DNA-binding domain and tested for interaction with HopZ1a$^{C216A}$ cloned as an in-frame fusion to the activation domain in the LexA yeast two-hybrid assay. As a positive control for interaction, we used HopF2$_{PtoDC3000}$ and its chaperone ShcF2$_{PtoDC3000}$ that are known to strongly interact[10].

*WAKL kinase domains do not trigger HR when co-expressed with HopZ1a*

To test if any of the WAKL kinase domains were capable of mediating ETI, we transiently co-expressed each WAKL kinase domain cloned as an in-frame fusion to a cYFP tag with HopZ1a in *N. benthamiana*. We did not observe an HR for any WAKLs when co-expressed with HopZ1a (Figure 5). As expected, we observed a strong HR when ZED1 and HopZ1a where co-expressed, and no HR when ZED1 and HopZ1a$^{C216A}$ were expressed (Figure 5)[11].

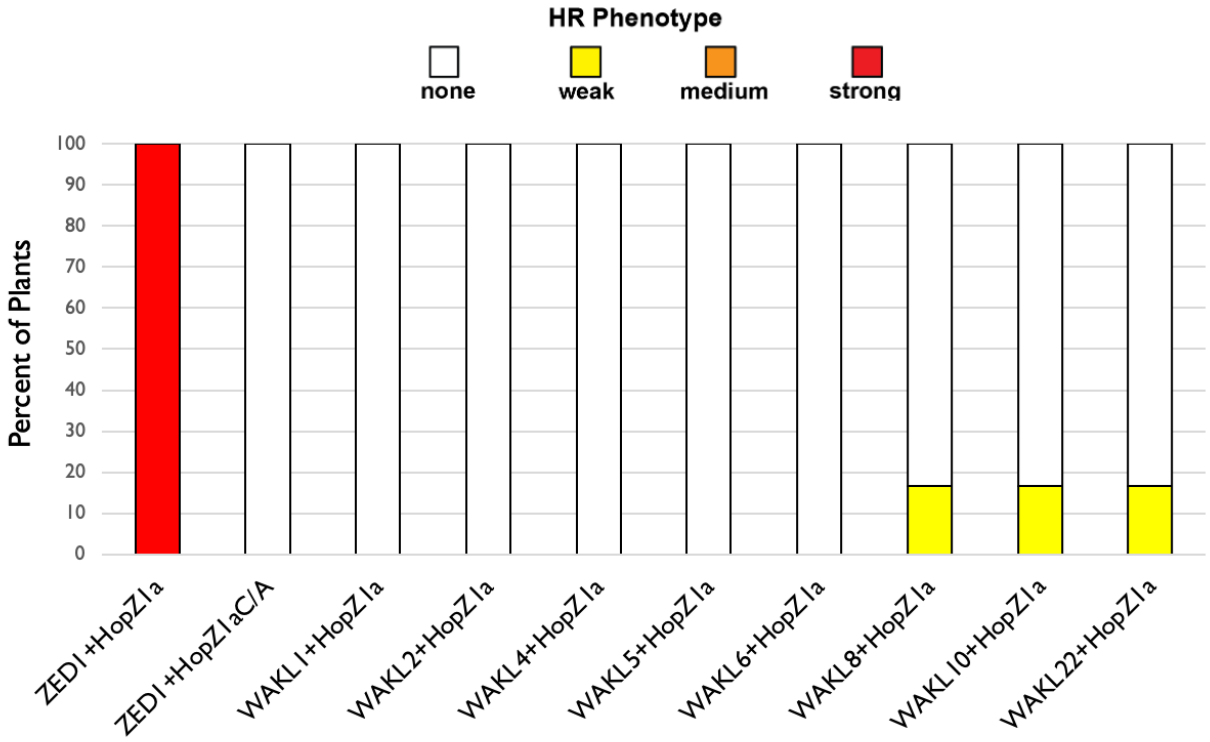



**Figure 5**. Kinase domains of WAKLs do not cause an HR when co-expressed with HopZ1a. Kinase domains of WAKLs were cloned into a dexamethasone-inducible expression vector and transiently co-expressed with HopZ1a in *N. benthamiana*. The HR was scored on a 4-point scale, 24 hours after dexamethasone induction (n=6 plants). ZED1 and HopZ1a were expressed as a positive control, and ZED1 and HopZ1a$^{C/A}$ were expressed as a negative control. Experiment was repeated once.

## Bibliography


1. Gong, Z. *et al.* The origin and evolution of a plant resistosome. *Plant Cell* **34**, 1600–1620 (2022).
2. Stephens, C., Hammond-Kosack, K. E. & Kanyuka, K. WAKsing plant immunity, waning diseases. *J. Exp. Bot.* **73**, 22–37 (2022).
3. Zhang, N., Pombo, M. A., Rosli, H. G. & Martin, G. B. Tomato Wall-Associated Kinase SlWak1 Depends on Fls2/Fls3 to Promote Apoplastic Immune Responses to Pseudomonas syringae. *Plant Physiol.* **183**, 1869–1882 (2020).
4. Verica, J. A. & He, Z. H. The cell wall-associated kinase (WAK) and WAK-like kinase gene family. *Plant Physiol.* **129**, 455–459 (2002).
5. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the *Pseudomonas syringae*type III effector HopZ1a. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18722–18727 (2013).
6. Diplock, N. *et al.* Utilising natural diversity of kinases to rationally engineer interactions with the angiosperm immune receptor ZAR1. *Plant. Cell Environ.* **46**, 2238–2254 (2023).
7. Waese, J. *et al.* ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *Plant Cell* **29**, 1806–1821 (2017).
8. Baudin, M. *et al.* A natural diversity screen in Arabidopsis thaliana reveals determinants for HopZ1a recognition in the ZAR1-ZED1 immune complex. *Plant. Cell Environ.* **44**, 629–644 (2021).

9. Huerta, A. I. *et al.* The WAK-like protein RFO1 acts as a sensor of the pectin methylation status in Arabidopsis cell walls to modulate root growth and defense. *Mol. Plant* **16**, 865–881 (2023).
10. Shan, L. *et al.* The HopPtoF locus of Pseudomonas syringae pv. tomato DC3000 encodes a type III chaperone and a cognate effector. *Mol. Plant-Microbe Interact.* **17**, 447–455 (2004).
11. Baudin, M., Hassan, J. A., Schreiber, K. J. & Lewis, J. D. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol.* **174**, 2038–2053 (2017).

## CONCLUSION

Here, we identify several Wall Associated Kinase-Like proteins which are targeted by the *P. syringae* effector HopZ1a. Connecting ZAR1 guardees with effector targets is a key step towards rational design of effector recognition through ZAR1.

**APPENDIX II:** Generation of a *N. benthamiana* ZAR1 knockout using CRISPR-Cas9

## PREFACE

The work presented in this appendix was a collaboration with Rachelle Punsalan and Jennifer D. Lewis. I performed the transient expression HR assays and genotyping with Rachelle Punsalan. I wrote the appendix with assistance from Jennifer D. Lewis. The work has been reproduced below with the permission of all authors.

Generation of a *N. benthamiana* ZAR1 knockout will further enable investigation of ZAR1 recognition pathways.

### Methods

Guide RNA sequences were designed using CRISPRdirect[1] and the *N. benthamiana* genome v1.0.1. Out of the two top scoring guide sequences, the guide closer to the 5' end of *N. benthamiana* ZAR1 (*NbZAR1*) was prioritized. The Guide RNA sequence (ACTTCTGCATGTTCTCACAG) was synthesized by fusing two oligonucleotides with complimentary overhanging sequences for ligation into the psgR-Cas9-At backbone following *Bbs*I restriction enzyme digest. An LR reaction was performed to transfer the psgR-Cas9 cassette into pEarleyGate303, and this vector was transformed into *A. tumefaciens* EHA105. Transgenic plants were generated at the UC Davis plant transformation center, using *N. benthamiana* seeds from a parent that produced an HR when ZED1 and HopZ1a were transiently expressed. Plant growth conditions and transient gene expression assays were performed as described in Chapter 3. PCR amplification for genotyping and sequencing was performed using GoTaq polymerase, and DNA was extracted following the Edwards method[2].

### Results

Transformation resulted in 15 individual T0 lines. To identify T1 individuals with homozygous mutations resulting in loss of *NbZAR1* function, we selected 6 independent lines T0 lines and grew 18 plants from each. We transiently expressed either HopZ1a and ZED1 or HopZ1a, ZED1, and *A. thaliana* ZAR1 (AtZAR1), or HopZ1a$^{C/A}$ and ZED1 in each T1 plant. We have previously shown that *NbZAR1*-silenced plants do not produce an HR when ZED1 and HopZ1a are co-expressed, while co-expressing ZED1, HopZ1a, and AtZAR1 in *NbZAR1*-silenced plants results in an HR[3]. To identify putative *nbzar1* plants, we looked for T1 individuals with no HR for HopZ1a and ZED1 or HopZ1a$^{C/A}$ and ZED1, but strong HR for ZED1, HopZ1a, and AtZAR1. This would indicate that plants were ETI-deficient for HopZ1a recognition through NbZAR1, but not ETI mediated through AtZAR1. Out of roughly 108 T1 plants screened, we identified 14 individuals with partial or complete loss of HR when HopZ1a and ZED1 were transiently co-expressed. We sequenced the 5' end of *NbZAR1* near the predicted cut site and found that 10 T1 individuals had an A insertion at the Cas9 cut site (following the 52$^{nd}$ nucleotide of the *NbZAR1* CDS), resulting in a premature stop codon following the 20$^{th}$ amino acid of NbZAR1 (Figure 1). In addition, 3 T1 individuals were heterozygous and specific modifications could not be identified through Sanger sequencing of *NbZAR1*. One of these heterozygous individuals was taken to the T2 generation, and T2 individuals were identified with a 6-nucleotide deletion in the *NbZAR1* CDS, resulting in deletion of the 19$^{th}$ and 20$^{th}$ amino acid of NbZAR1 (Figure 1). To attempt to remove the *Cas9* gene, we allowed T1 plants to self and genotyped 22, *zar1* T2 individuals for loss of the *Cas9* gene. All T2 individuals with

homozygous mutations in *NbZAR1* maintained the *Cas9* gene. We did not notice any obvious morphological, developmental, or fitness phenotypes in the *zar1* lines relative to WT *N. benthamiana*.

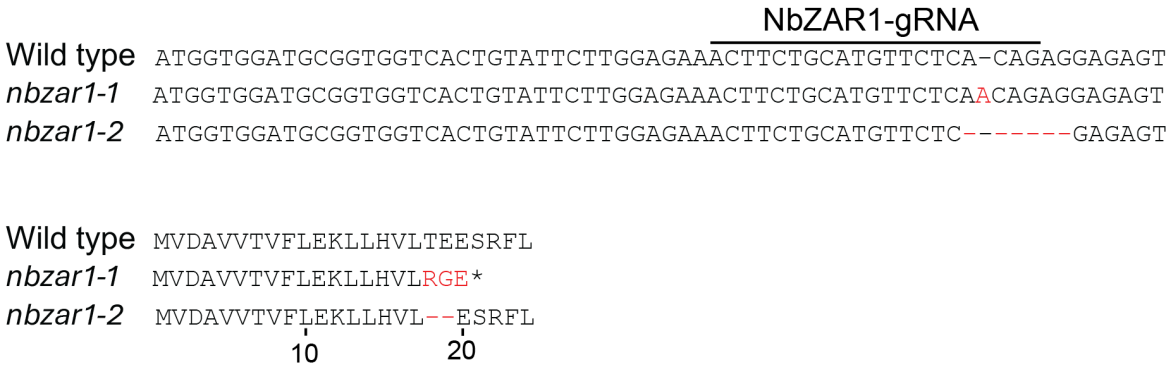



**Figure 1.** DNA and amino acid sequence of *nbzar1-1* and *nbzar1-2* relative to wild-type. Red text indicates changes in *nbzar1-1* sequence relative to wild-type.

**Bibliography**


1. Naito, Y., Hino, K., Bono, H. & Ui-Tei, K. CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* **31**, 1120–1123 (2015).
2. Edwards, K., Johnstone, C. & Thompson, C. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**, 1349–1349 (1991).
3. Baudin, M., Hassan, J. A., Schreiber, K. J. & Lewis, J. D. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. *Plant Physiol.* **174**, 2038–2053 (2017).


## CONCLUSION

Here, we generate a *N. benthamiana* ZAR1 knockout which will further enable investigation of ZAR1 recognition systems.