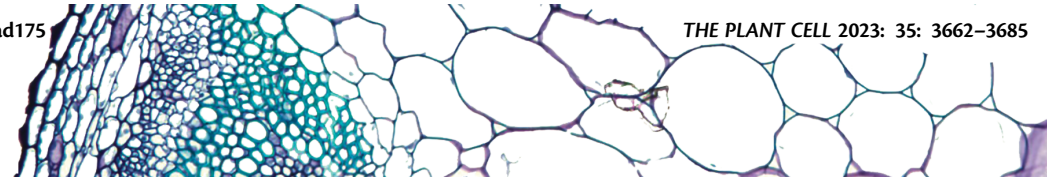
https://doi.org/10.1093/plcell/koad175 THE PLANT CELL 2023: 35: 3662–3685



Large-Scale Biology

Jurassic NLR: Conserved and dynamic evolutionary features of the atypically ancient immune receptor ZAR1

Hiroaki Adachi [1,2,3] Toshiyuki Sakai [1,2] Jiorgos Kourelis [1] Hsuan Pai [1] Jose L. Gonzalez Hernandez [4] Yoshinori Utsumi [5] Motoaki Seki [5,6,7] Abbas Maqbool [1] and Sophien Kamoun [1,*]

1 The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, Norwich NR4 7UH, UK
2 Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University, Mozume, Muko, Kyoto 617-0001, Japan
3 PRESTO, Japan Science and Technology Agency, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan
4 Agronomy, Horticulture and Plant Sciences Department, South Dakota State University, Brookings, SD 57007, USA
5 Plant Genomic Network Research Team, RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
6 Plant Epigenome Regulation Laboratory, RIKEN Cluster for Pioneering Research, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
7 Kihara Institute for Biological Research, Yokohama City University, 641-12 Maioka-cho, Totsuka-ku, Yokohama, Kanagawa 244-0813, Japan

*Author for correspondence: sophien.kamoun@tsl.ac.uk
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) is: Sophien Kamoun (sophien.kamoun@tsl.ac.uk).

Abstract

Plant nucleotide-binding leucine-rich repeat (NLR) immune receptors generally exhibit hallmarks of rapid evolution, even at the intraspecific level. We used iterative sequence similarity searches coupled with phylogenetic analyses to reconstruct the evolutionary history of HOPZ-ACTIVATED RESISTANCE1 (ZAR1), an atypically conserved NLR that traces its origin to early flowering plant lineages ~220 to 150 million yrs ago (Jurassic period). We discovered 120 ZAR1 orthologs in 88 species, including the monocot *Colocasia esculenta*, the magnoliid *Cinnamomum micranthum*, and most eudicots, notably the Ranunculales species *Aquilegia coerulea*, which is outside the core eudicots. Ortholog sequence analyses revealed highly conserved features of ZAR1, including regions for pathogen effector recognition and cell death activation. We functionally reconstructed the cell death activity of ZAR1 and its partner receptor-like cytoplasmic kinase (RLCK) from distantly related plant species, experimentally validating the hypothesis that ZAR1 evolved to partner with RLCKs early in its evolution. In addition, ZAR1 acquired novel molecular features. In cassava (*Manihot esculenta*) and cotton (*Gossypium* spp.), ZAR1 carries a C-terminal thioredoxin-like domain, and in several taxa, *ZAR1* duplicated into 2 paralog families, which underwent distinct evolutionary paths. *ZAR1* stands out among angiosperm NLR genes for having experienced relatively limited duplication and expansion throughout its deep evolutionary history. Nonetheless, ZAR1 also gave rise to noncanonical NLRs with integrated domains and degenerated molecular features.

Introduction

Plant immune receptors, often encoded by disease resistance (*R*) genes, detect invading pathogens and activate innate immune responses that can limit infection (Jones and Dangl 2006). A major class of immune receptors is formed by intracellular proteins of the nucleotide-binding leucine-rich repeat (NLR) family (Dodds and Rathjen 2010; Jones et al. 2016; Kourelis and van der Hoorn 2018). NLRs detect host-translocated pathogen effectors either by directly binding them or by indirectly monitoring proteins known as guardees or decoys.

NLRs are arguably the most diverse protein family in flowering plants (angiosperms) with many species having large (>100) and diverse repertoires of NLR genes in their genomes (Shao et al. 2016; Baggs et al. 2017; Kourelis et al. 2021). NLR genes typically exhibit hallmarks of rapid evolution even at the intraspecific level (Van de Weyer et al. 2019; Lee and

Received January 17, 2023. Accepted June 02, 2023. Advance access publication July 19, 2023

IN A NUTSHELL

Background: In plants, nucleotide-binding leucine-rich repeat (NLR) immune receptors generally exhibit hallmarks of rapid evolution even at the intraspecific level. NLRs evolve primarily through a birth-and-death process: new NLRs emerge by recurrent cycles of gene duplication and loss—some genes are maintained in the genome and acquire new pathogen detection specificities, whereas others are deleted or become nonfunctional through the accumulation of deleterious mutations. Such dynamic patterns of evolution enable the NLR immune system to keep up with fast-evolving effector repertoires of pathogenic microbes. Unlike typical NLRs, HOPZ-ACTIVATED RESISTANCE1 (ZAR1) is conserved across angiosperms.

Question: Can we use a molecular evolution framework to determine the critical features of a conserved plant NLR?

Findings: We performed iterative sequence similarity searches coupled with phylogenetic analyses to reconstruct the evolutionary history of ZAR1. ZAR1 is an atypically conserved NLR that traces its origin to early flowering plant lineages ~220 to 150 million yrs ago (Jurassic period). Ortholog sequence analyses revealed highly conserved features of ZAR1, including regions for pathogen recognition and immune activation. We functionally reconstructed the immune activity of ZAR1 and its host partner receptor-like cytoplasmic kinases (RLCKs) from distantly related plant species, supporting the hypothesis that ZAR1 has evolved to partner with RLCKs early in its evolution. ZAR1 stands out among angiosperm NLRs for having experienced relatively limited gene duplication and expansion throughout its deep evolutionary history.

Next steps: Further comparative analyses, combining molecular evolution and structural biology of plant and animal NLR systems will yield experimentally testable hypotheses for NLR research.

Chae 2020; Prigozhin and Krasileva 2020). Towards the end of the 20th century, Michelmore and Meyers (1998) proposed that NLRs evolve primarily through the birth-and-death process. In this model, new NLRs emerge by recurrent cycles of gene duplication and loss—some genes are maintained in the genome and acquire new pathogen detection specificities, whereas others are deleted or become nonfunctional through the accumulation of deleterious mutations. Such dynamic patterns of evolution enable the NLR immune system to keep up with fast-evolving effector repertoires of pathogenic microbes.

However, as noted over 20 yrs ago by Michelmore and Meyers (1998), a subset of NLR proteins is slow evolving and has remained fairly conserved throughout evolutionary time (Wu et al. 2017; Stam et al. 2019). These "high-fidelity" NLRs (per Lee and Chae 2020) offer unique opportunities for comparative analyses, providing a molecular evolution framework to reconstruct key transitions and reveal functionally critical biochemical features (Delaux et al. 2019). Nonetheless, comprehensive evolutionary reconstructions of conserved NLR proteins remain limited despite the availability of a large number of plant genomes across the breadth of plant phylogeny. One of the reasons is that the great majority of NLRs lack clear-cut orthologs across divergent plant taxa. Here, we address this gap in knowledge by investigating the macroevolution of ZAR1 (HOPZ-ACTIVATED RESISTANCE1), an atypically ancient NLR, and asking fundamental questions about the conservation and diversification of this immune receptor throughout its deep evolutionary history.

NLRs generally function in nonself perception and innate immunity in plants and animals (Jones et al. 2016; Uehling et al. 2017). In the broadest biochemical definition, plant NLRs share a multidomain architecture typically consisting of a NB-ARC (nucleotide-binding domain shared with APAF-1, various R-proteins, and CED-4) followed by a leucine-rich repeat (LRR) domain. Angiosperm NLRs form several major monophyletic groups with distinct N-terminal domain fusions (Shao et al. 2016; Kourelis et al. 2021). These include the subclades TIR-NLR with the Toll/interleukin-1 receptor (TIR) domain, CC-NLR with the Rx-type coiled-coil (CC) domain, CC_R-NLR with the RPW8-type CC (CC_R) domain (Tamborski and Krasileva 2020), and the more recently defined CC_{G10}-NLR with a distinct type of CC (CC_{G10}) (Lee H-Y et al. 2020). In addition to the canonical tripartite domain architecture, up to 10% of NLRs carry unconventional "integrated" domains. Integrated domains (IDs) are thought to generally function as decoys to bait pathogen effectors and enable pathogen detection (Cesari et al. 2014; Wu et al. 2015; Sarris et al. 2016; Kourelis and van der Hoorn 2018). They include dozens of different modules indicating that novel domain acquisitions have repeatedly taken place throughout the evolution of plant NLRs (Kroj et al. 2016; Sarris et al. 2016). To date, over 400 NLRs from 31 genera in 11 orders of flowering plants have been experimentally validated as reported in the RefPlantNLR reference dataset (Kourelis et al. 2021). Several of these NLRs are coded by *R* genes that function against economically important pathogens and contribute to sustainable agriculture (Dangl et al. 2013).

In recent years, the research community has gained a better understanding of the structure/function relationships of plant NLRs and the immune receptor circuitry they form (Wu et al. 2018; Adachi et al. 2019a; Burdett et al. 2019; Jubic et al. 2019;

Bayless and Nishimura 2020; Feehan et al. 2020; Mermigka et al. 2020; Wang and Chai 2020; Xiong et al. 2020; Zhou and Zhang 2020). Some NLRs, such as ZAR1, form a single functional unit that carries both pathogen sensing and immune signaling activities in a single protein (termed "singleton NLR" per Adachi et al. 2019a). Other NLRs function together in pairs or more complex networks, where connected NLRs have functionally specialized into sensor NLRs dedicated to pathogen detection or helper NLRs that are required for sensor NLRs to initiate immune signaling (Feehan et al. 2020). Paired and networked NLRs are thought to have evolved from multifunctional ancestral receptors through asymmetrical evolution (Adachi et al. 2019a, 2019b). As a result of their direct coevolution with pathogens, NLR sensors tend to diversify faster than helpers and can be dramatically expanded in some plant taxa (Wu et al. 2017; Stam et al. 2019). For instance, sensor NLRs often exhibit noncanonical biochemical features, such as degenerated functional motifs and unconventional domain integrations (Adachi et al. 2019b; Seong et al. 2020).

The elucidation of plant NLR structures by cryo-electron microscopy has significantly advanced our understanding of the biochemical events associated with the activation of these immune receptors (Wang et al., 2019a, 2019b; Ma et al. 2020; Martin et al. 2020). The CC-NLR ZAR1, the TIR-NLRs RPP1 and Roq1 oligomerize upon activation into a multimeric complex known as the resistosome. In the case of ZAR1, recognition of bacterial effectors occurs through its partner receptor-like cytoplasmic kinases (RLCKs), which are encoded by genes that occur in a genomic cluster; this cluster encodes multiple RLCK-type pseudokinases that vary depending on the pathogen effector and host plant (Lewis et al. 2013; Wang et al., 2015; Seto et al. 2017; Schultink et al. 2019; Laflamme et al. 2020). Activation of ZAR1 induces conformational changes in the nucleotide-binding domain resulting in ADP release, dATP/ATP binding, and pentamerization of the ZAR1–RLCK complex into the resistosome. The ZAR1 resistosome exposes a funnel-shaped structure formed by the N-terminal α1 helices, which translocates into the plasma membrane, and the resistosome itself acts as a Ca^{2+} channel (Wang et al., 2019b; Bi et al. 2021). The ZAR1 N-terminal α1 helix matches the MADA consensus sequence motif that is functionally conserved in ∼20% of CC-NLRs including NLRs from dicot and monocot plant species (Adachi et al. 2019b). This suggests that the biochemical "death switch" mechanism of the ZAR1 resistosome may apply to a significant fraction of CC-NLRs. Interestingly, unlike singleton and helper CC-NLRs, sensor CC-NLRs often carry degenerated MADA α1 helix motifs and/or N-terminal domain integrations, which would preclude their capacity to trigger cell death according to the ZAR1 model (Adachi et al. 2019b; Seong et al. 2020).

Comparative sequence analyses based on a robust evolutionary framework can yield insights into molecular mechanisms and help generate experimentally testable hypotheses. ZAR1 was previously reported to be conserved across multiple dicot plant species but whether it occurs in other angiosperms hasn't been systematically studied (Lewis et al. 2010; Baudin et al. 2017; Schultink et al. 2019; Harant et al. 2022). Here, we used a phylogenomic approach to investigate the molecular evolution of ZAR1 across flowering plants (angiosperms). We discovered 120 ZAR1 orthologs in 88 species, including monocot, magnoliid, and eudicot species indicating that ZAR1 is an atypically conserved CC-NLR that traces its origin to early angiosperm lineages ∼220 to 150 million yrs ago (Jurassic period). We took advantage of this large collection of orthologs to identify highly conserved features of ZAR1, revealing regions for effector recognition, intramolecular interactions, and cell death activation. We showed that the cell death activity of ZAR1 from distantly related plant species can be dependent of its partner RLCKs, therefore experimentally validating the hypothesis that ZAR1 has evolved to be a partner with RLCKs early in its evolution. Throughout its evolution, ZAR1 also acquired features, including the C-terminal integration of a thioredoxin-like domain and duplication into 2 paralog families ZAR1-SUB and ZAR1-CIN. Members of the ZAR1-SUB paralog family have highly diversified in eudicots and often lack conserved ZAR1 features. We conclude that ZAR1 has experienced relatively limited gene duplication and expansion throughout its deep evolutionary history but still did give rise to noncanonical NLR proteins with IDs and degenerated molecular features.

Results

ZAR1 is widely conserved across angiosperms

To determine the distribution of ZAR1 across plant species, we applied a computational pipeline based on iterated BLAST searches of plant genome and protein databases (Fig. 1A). These comprehensive searches were seeded with previously identified ZAR1 sequences from Arabidopsis (*Arabidopsis thaliana*), *Nicotiana benthamiana*, tomato (*Solanum lycopersicum*), sugar beet (*Beta vulgaris*), and cassava (*Manihot esculenta*) (Baudin et al. 2017; Schultink et al. 2019; Harant et al. 2022). We also performed iterated phylogenetic analyses using the NB-ARC domain of the harvested ZAR1-like sequences and obtained a well-supported clade that includes the previously reported ZAR1 sequences, as well as new clade members from more distantly related plant species, notably *Colocasia esculenta* (taro, Alismatales), *Cinnamomum micranthum* (Syn. *Cinnamomum kanehirae*, stout camphor, Magnoliidae), and *Aquilegia coerulea* (columbine, Ranunculales; Supplemental Data Set 1).

In total, we identified 120 *ZAR1* genes from 88 angiosperm species that tightly clustered in the ZAR1 phylogenetic clade (Fig. 1B and Supplemental Data Set 1). Among the 120 genes, 108 code for canonical CC-NLR proteins with 52.0% to 97.0% similarity to Arabidopsis ZAR1, whereas another 9 carry the 3 major domains of CC-NLR proteins but have a C-terminal integrated domain (ZAR1-ID, see below). The remaining 3 genes code for 2 truncated NLRs and a potentially misannotated coding sequence due to a gap in the genome sequence. In summary, we propose that the identified clade

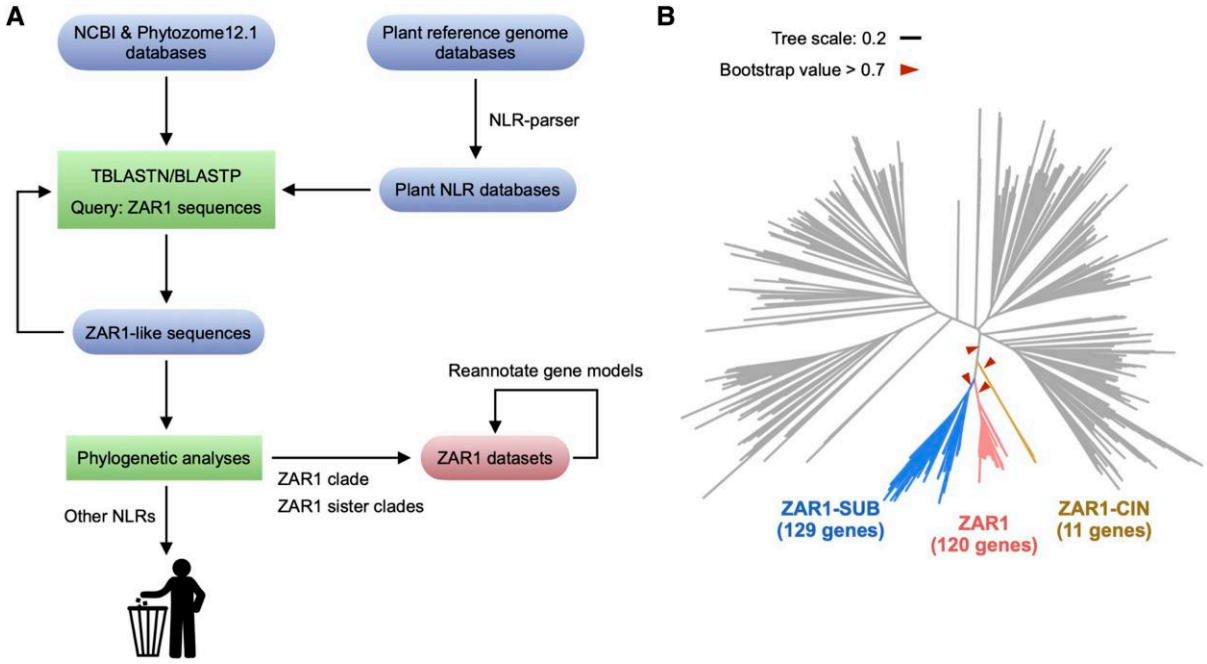


Figure 1. Comparative sequence analyses identify and classify ZAR1 sequences from angiosperms. **A)** Workflow for computational analyses in searching for *ZAR1* orthologs. We performed TBLASTN/BLASTP searches and subsequent phylogenetic analyses to identify *ZAR1* ortholog genes from plant genome/proteome datasets. **B)** ZAR1 forms a clade with 2 closely related sister subclades. The phylogenetic tree was generated in MEGA7 by the neighbor-joining method using NB-ARC domain sequences of ZAR1-like proteins identified from the prior BLAST searches and 1,019 NLRs identified from 6 representative plant species, taro, stout camphor, columbine, tomato, sugar beet, and Arabidopsis. Red arrowheads indicate bootstrap support > 0.7 and is shown for the relevant nodes. The scale bar indicates the evolutionary distance in amino acid substitution per site.

consists of ZAR1 orthologs from a diversity of angiosperm species. Our analyses of ZAR1-like sequences also revealed 2 well-supported sister clades of the ZAR1 ortholog clade (Fig. 1B). We named these subclades ZAR1-SUB and ZAR1-CIN (referred to as ZAR1-sis and ZAR1-basal in Gong et al. (2022), respectively), and we describe them in more detail below.

We have recently proposed that ZAR1 is the most conserved CC-NLR between rosid and asterid plants (Harant et al. 2022). To further evaluate ZAR1 conservation relative to other CC-NLRs across angiosperms, we used a phylogenetic tree of 1,475 NLRs from the monocot taro, the magnoliid stout camphor, and 6 eudicot species (columbine, Arabidopsis, cassava, sugar beet, tomato, and *N. benthamiana*) to calculate the phylogenetic (patristic) distance between each of the 49 Arabidopsis CC-NLRs and their closest neighbor from each of the other plant species. As shown in Harant et al. (2022), ZAR1 stands out for having the shortest phylogenetic distance to its orthologs relative to other CC-NLRs in this diverse angiosperm species set (Supplemental Fig. S1). A similar analysis where we plotted the phylogenetic distance between each of the 159 *N. benthamiana* CC-NLRs to their closest neighbor from the other species also revealed ZAR1 as displaying the shortest patristic distance across all examined species (Supplemental Fig. S2). These analyses revealed that ZAR1 is possibly the most widely conserved CC-NLR in flowering plants (angiosperms).

Phylogenetic distribution of *ZAR1* in angiosperms

Although *ZAR1* is distributed across a wide range of angiosperms, we noted particular patterns in its phylogenetic distribution. Supplemental Data Set 1 describes the gene identifiers and other features of ZAR1 orthologs sorted based on the phylogenetic clades reported by Smith and Brown (2018). Sixty-eight of the 88 plant species have a single copy of *ZAR1* whereas 20 species have 2 or more copies (Supplemental Data Set 2). *ZAR1* is primarily a eudicot gene, but we identified 3 *ZAR1* orthologs outside the eudicots, 2 in the monocot taro, and another 1 in the magnoliid stout camphor. We failed to detect *ZAR1* orthologs in 39 species among the 127 species we examined (Supplemental Data Set 1). Except for taro, *ZAR1* is missing in monocot species (17 examined), including in the well-studied *Hordeum vulgare* (barley), *Oryza sativa* (rice), *Triticum aestivum* (wheat), and *Zea mays* (maize). *ZAR1* is also missing in all examined species of the eudicot Fabales, Cucurbitales, Apiales, and Asterales. However, we found a *ZAR1* ortholog in columbine (a eudicot outside the monophyletic clade of core eudicots) and *ZAR1* is widespread in other eudicots, including in 63 rosid, 4 Caryophyllales, and 18 asterid species.

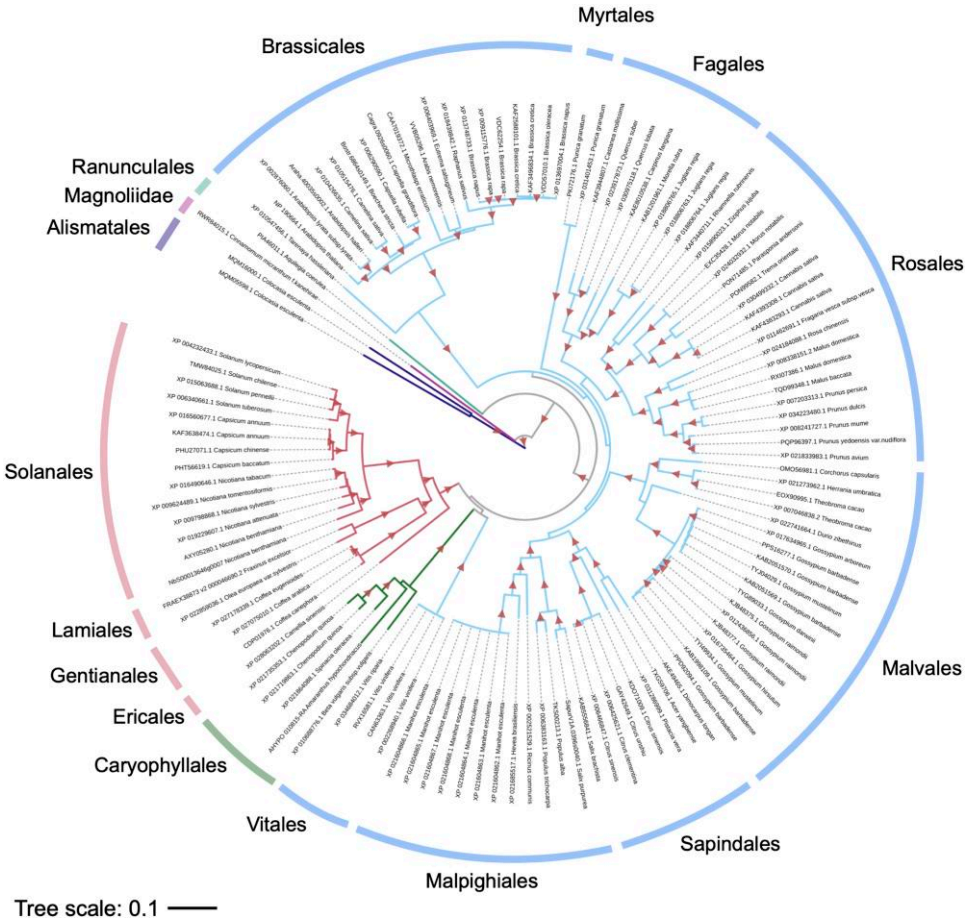


Figure 2. The *ZAR1* gene is distributed across angiosperms. The phylogenetic tree was generated in MEGA7 by the neighbor-joining method using full-length amino acid sequences of 120 ZAR1 orthologs identified in Fig. 1. Red triangles indicate bootstrap support > 0.7. The scale bar indicates the evolutionary distance in amino acid substitution per site.

ZAR1 is an ancient Jurassic gene that predates the split between monocots, magnoliids, and eudicots

The overall conservation of the 120 ZAR1 orthologs enabled us to perform phylogenetic analyses using the full-length protein sequence and not just the NB-ARC domain as generally done with NLRs (Fig. 2 and Supplemental Fig. S3). These analyses yielded a robust ZAR1 phylogenetic tree with well-supported branches that generally mirrored established phylogenetic relationships between the examined plant species (Smith and Brown 2018; Chaw et al. 2019). For example, the ZAR1 tree matched a previously published species tree of angiosperms based on 211 single-copy core ortholog genes (Chaw et al. 2019). We conclude that the origin of the *ZAR1* gene predates the split between monocots, magnoliids, and eudicots and its evolution traced species divergence ever since. We postulate that ZAR1 probably emerged in the Jurassic era ~220 to 150 million yrs ago (Mya) based on the species divergence time estimate of Chaw et al. (2019) and consistent with the latest fossil evidence for the emergence of flowering plants (Fu et al. 2018; Cui et al. 2022).

ZAR1 is a genetic singleton in a locus that exhibits gene colinearity across eudicot species

NLR genes are often clustered in loci that are thought to accelerate sequence diversification and evolution (Michelmore and Meyers 1998; Lee and Chae 2020). We examined the genetic context of ZAR1 genes using available genome assemblies of taro, stout camphor, columbine, Arabidopsis, cassava, sugar beet, tomato, and *N. benthamiana*. The *ZAR1* locus is generally devoid of other NLR genes as the closest NLR is found in the Arabidopsis genome 183 kb away from *ZAR1* (Supplemental Data Set 3). We conclude that *ZAR1* has probably remained a genetic singleton NLR gene throughout its evolutionary history in angiosperms.

Next, we examined the *ZAR1* locus for gene colinearity across the examined species. We noted a limited degree of gene colinearity between Arabidopsis vs. cassava, cassava vs. tomato, and tomato vs. *N. benthamiana* (Supplemental Fig. S4). Flanking conserved genes include the ATPase and protein kinase genes that are present at the *ZAR1* locus in both rosid and asterid eudicots. In contrast, we didn't observe conserved gene blocks at the *ZAR1* locus of taro, stout

camphor, and columbine, indicating that this locus is divergent in these species. Overall, although limited, the observed gene colinearity in eudicots is consistent with the conclusion that *ZAR1* is a genetic singleton with an ancient origin.

ZAR1 orthologs carry sequence motifs known to be required for Arabidopsis ZAR1 resistosome function

The overall sequence conservation and deep evolutionary origin of ZAR1 orthologs combined with the detailed knowledge of the ZAR1 structure and function provide a unique opportunity to explore the evolutionary dynamics of this ancient immune receptor in a manner that cannot be applied to more rapidly evolving NLRs. We used MEME (Multiple EM for Motif Elicitation; Bailey and Elkan 1994) to search for conserved sequence patterns among the 117 ZAR1 orthologs (ZAR1 and ZAR1-ID) that encode full-length CC-NLR proteins. This analysis revealed several conserved sequence motifs that span across the ZAR1 orthologs (range of protein lengths: 753–1,132 amino acids; Fig. 3A and Supplemental Table S1). In Fig. 3A, we described the major 5 sequence motifs or interfaces known to be required for Arabidopsis ZAR1 function that are conserved across ZAR1 orthologs.

Effector recognition by ZAR1 occurs indirectly via binding to RLCKs through the LRR domain. Key residues in the Arabidopsis ZAR1–RLCK interfaces are highly conserved among ZAR1 orthologs and were identified by MEME as conserved sequence patterns (Fig. 3A). Valine (V) 544, histidine (H) 597, glycine (G) 645, proline (P) 816, tryptophan (W) 825, and phenylalanine (F) 839 in the Arabidopsis ZAR1 LRR domain were validated by mutagenesis as important residues for RLCK binding whereas isoleucine (I) 600 was not essential (Wang et al., 2015; Baudin et al. 2017; Wang et al., 2019a; Hu et al. 2020). In the 117 ZAR1 orthologs, V544, H597, G645, P816, W825, and F839 are conserved in 88% to 100% of the proteins compared to only 63% for I600.

After effector recognition, Arabidopsis ZAR1 undergoes conformational changes from the monomeric inactive form to the oligomeric active state. This is mediated by ADP release from the NB-ARC domain and subsequent ATP binding, which triggers further structural remodeling in ZAR1 leading to the formation of the activated pentameric resistosome (Wang et al., 2019b). NB-ARC sequences that coordinate binding and hydrolysis of dATP, namely, P-loop and MHD motifs, are highly conserved across ZAR1 orthologs (Fig. 3A). Histidine (H) 488 and lysine (K) 195, located in the ADP/ATP-binding pocket (Wang et al., 2019a, 2019b), are invariant in all 117 orthologs. In addition, 3 NB-ARC residues, W150, S152, and V154, known to form the NBD–NBD oligomerization interface for resistosome formation (Wang et al., 2019b; Hu et al. 2020), are present in 82% to 97% of the ZAR1 orthologs and were also part of a MEME motif (Fig. 3A).

The N-terminal CC domain of Arabidopsis ZAR1 mediates cell death signaling thorough the N-terminal α1 helix/MADA motif that becomes exposed in activated ZAR1 resistosome to form a funnel-like structure (Baudin et al. 2017, 2019; Wang et al., 2019b; Adachi et al. 2019b). We detected an N-terminal MEME motif that matches the α1 helix/MADA motif (Fig. 3A). We also used the HMMER software (Eddy 1998) to query the ZAR1 orthologs with a previously reported MADA motif–hidden Markov model (HMM; Adachi et al. 2019b). This HMMER search detected a MADA-like sequence at the N-terminus of all 117 ZAR1 orthologs (Supplemental Data Set 1).

Taken together, based on the conserved motifs depicted in Fig. 3A, we propose that angiosperm ZAR1 orthologs share the main functional features of Arabidopsis ZAR1: (i) effector recognition via RLCK binding, (ii) remodeling of intramolecular interactions via ADP/ATP switch, (iii) oligomerization via the NBD–NBD interface, and (iv) α1 helix/MADA motif–mediated activation of hypersensitive cell death.

ZAR1 resistosome displays conserved surfaces on RLCK binding sites and the inner glutamate ring

To identify additional conserved and variable features in ZAR1 orthologs, we used ConSurf (Ashkenazy et al. 2016) to calculate a conservation score for each amino acid and generate a diversity barcode for ZAR1 orthologs (Fig. 3B). We then used the cryo-EM structures of Arabidopsis ZAR1 to determine how the ConSurf score maps onto the 3D structures (Figs. 3, C and D, and 4). First, we found 5 major variable surfaces (VS1 to VS5) on the inactive ZAR1 monomer structure (Fig. 3, C and D), as depicted in the ZAR1 diversity barcode (Fig. 3B). VS1 comprises α2/α4 helices and a loop between α3 and α4 helices of the CC domain. VS2 and VS3 corresponds to α1/α2 helices of NBD and a loop between α2 and α3 helices of HD1, respectively. VS4 comprises a loop between WHD and LRR and first 3 helices of the LRR domain. VS5 is mainly derived from the last 3 helices of the LRR domain and the loops between these helices (Fig. 3, B and D).

Next, we examined highly conserved surfaces on inactive and active ZAR1 structures (Fig. 4, A and B). Consistent with the MEME analyses, we confirmed that highly conserved surfaces match to the RLCK binding interfaces (Fig. 4, A and B). We also confirmed that the N-terminal α1 helix/MADA motif is conserved on the resistosome surfaces, although the first 4 N-terminal amino acids are missing from the N terminus of the active ZAR1 cryo-EM structures (Fig. 4, B and C). We also noted sequence conservation at the glutamate rings (comprised of E11, E18, E130, and E134) inside the Arabidopsis ZAR1 resistosome (Supplemental Fig. S5). Glutamic acid (E) 11 is conserved in 94% of ZAR1 orthologs, whereas only 3% to 18% retain E18, E130, and E134 in the same positions as Arabidopsis ZAR1. Interestingly, mutation of E11 to alanine (A) impaired Arabidopsis ZAR1–mediated cell death but the E18A, E130A, and E134A mutants were capable of inducing cell death (Bi et al. 2021). Furthermore, the E11A mutation impaired Ca^{2+} channel activity of the ZAR1 resistosome in vitro and in vivo (Bi et al. 2021). Therefore, our motif and structure analyses suggest that RLCK-mediated effector recognition and E11-dependent Ca^{2+} influx are key functional features conserved across the great majority of ZAR1 orthologs.

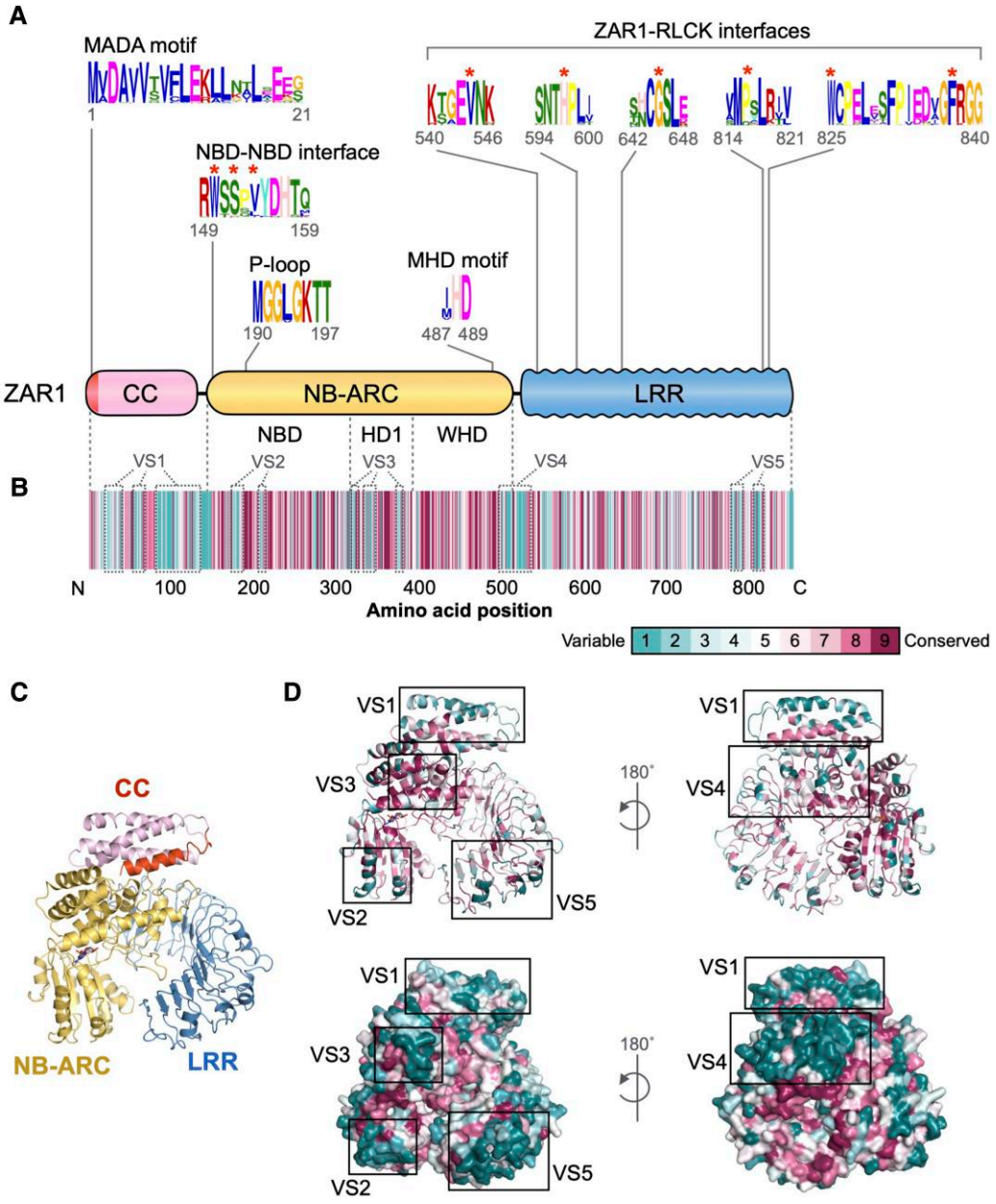


Figure 3. ZAR1 orthologs carry conserved sequence patterns required for Arabidopsis ZAR1 resistosome function. **A)** Schematic representation of the Arabidopsis ZAR1 protein highlighting the position of conserved sequence patterns across ZAR1 orthologs. Consensus sequence patterns were identified by MEME using 117 ZAR1 ortholog sequences. Raw MEME motifs are listed in Supplemental Table S1. Red asterisks indicate residues functionally validated in Arabidopsis ZAR1 for NBD–NBD and ZAR1–RLCK interfaces. **B)** Conservation and variation of each amino acid among ZAR1 orthologs across angiosperms. Amino acid alignment of 117 ZAR1 orthologs was used for conservation score calculation via the ConSurf server (https://consurf.tau.ac.il). The conservation scores are mapped onto each amino acid position in Arabidopsis ZAR1 (NP_190664.1). **C, D)** Distribution of the ConSurf conservation score on the Arabidopsis ZAR1 structure. The inactive ZAR1 monomer is illustrated in cartoon representation with domain architecture **C)** and conservation score **D)**. Major 5 variable surfaces (VS1 to VS5) on the inactive ZAR1 monomer structure are described in gray dot or black boxes in panel B or D, respectively.

ZAR1 interaction sites are conserved in ZED1-related kinase (ZRK) family proteins across distantly related plant species

We endeavored to experimentally test the hypothesis that ZAR1 ortholog proteins across angiosperm species require RLCKs to activate their molecular switch. First, we searched for RLCK XII-2 subfamily genes in the distantly related plant species, taro, stout camphor, and columbine. The BLAST searches of protein databases were seeded with previously identified RLCK ZED1-related kinase (ZRK) sequences from Arabidopsis and *N. benthamiana* (Lewis et al. 2013; Schultink et al. 2019). We also performed iterated phylogenetic analyses

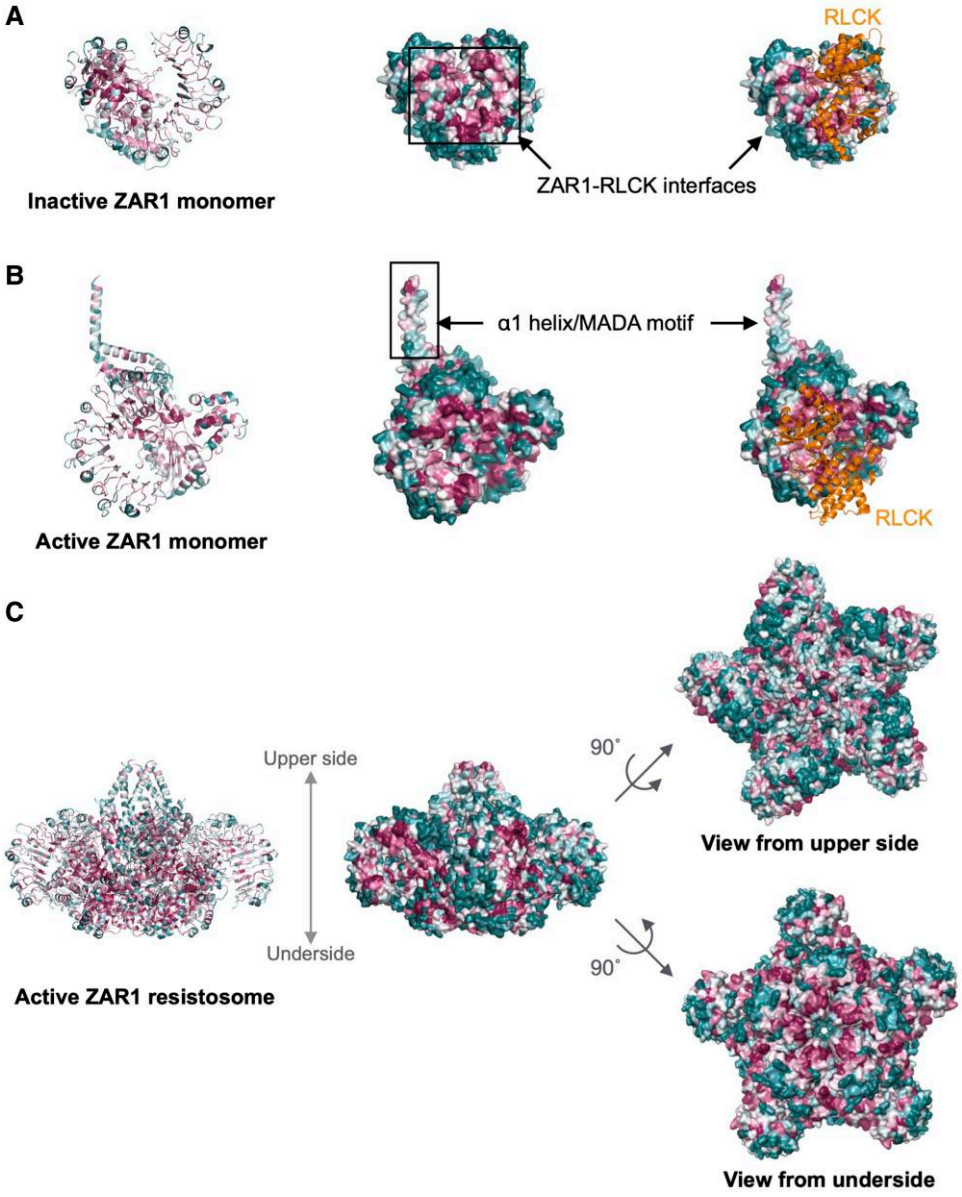


Figure 4. ZAR1 orthologs across angiosperms display multiple conserved surfaces on the resistosome structure. Distribution of the ConSurf conservation score was visualized on the inactive monomer **A)**, active monomer **B)**, and resistosome **C)** structures of Arabidopsis ZAR1. Each structure and cartoon representation are illustrated based on the conservation score shown in Fig. 3.

using the kinase domain of the harvested ZRK-like sequences and obtained a well-supported clade that includes previously reported ZRK from Arabidopsis (ZRK1∼7, 10∼15) and *N. benthamiana* (JIM2) as well as additional clade members from taro, stout camphor, and columbine (Fig. 5A). In total, we identified 21 *ZRK* genes in these species, which include 1 *ZRK* gene (*CeZRK1*) from taro, 15 *ZRK* genes (*CmZRK1∼15*) from stout camphor, and 5 *ZRK* genes (*AcZRK1∼5*) from columbine (Fig. 5A; Supplemental Data Set 4).

Remarkably, similar to Arabidopsis *ZRKs* (Lewis et al. 2013), a number of the identified *ZRKs* are located in genomic clusters. 13 *ZRK* genes in stout camphor and 4 *ZRK* genes in columbine form gene clusters on scaffold QPKB01000003.1 and contig KZ305039.1, respectively (Fig. 5B). All of the identified *ZRK* genes are located on a different scaffold or contig to the *ZAR1* gene in taro, stout camphor, and columbine, whereas Arabidopsis *ZAR1* and the 9 *ZRK* genes occur on the same chromosome (Supplemental Data Set 4).

The 21 *ZRK* genes code for proteins of 277 to 452 amino acids, similar to Arabidopsis and *N. benthamiana ZRKs*, which code for 269 to 396 amino acid proteins (Supplemental File 1). ZRK family proteins from taro, stout camphor, and columbine show 20.8% to 42.2% similarity to Arabidopsis RKS1/ZRK1 (Supplemental File 2). Although the sequence similarity is low across the ZRK proteins in angiosperms, ZAR1 interaction sites are highly conserved in the ZRKs (Supplemental

Fig. S6; Wang et al., 2019a; Hu et al. 2020). Notably, functionally validated residues for ZAR1–RLCK interactions (G27 and leucine 31 [L31] in Arabidopsis RKS1/ZRK1; G29 and aspartic acid [D] 231 in Arabidopsis ZED1/ZRK5) are conserved in 81% to 100% of the 21 ZRKs. Moreover, 90% of the 21 ZRKs have a hydrophobic V or I residue at the same position to V35 in Arabidopsis RKS1/ZRK1 (corresponding to I24 in Arabidopsis ZED1/ZRK5). This sequence conservation supports our hypothesis that ZRK family proteins function together with ZAR1 across distantly related plant species.

Heterologous expression of *ZAR1* and *ZRK* orthologs from flowering plant species in *N. benthamiana*

To validate functional connections between ZAR1 orthologs and their partner ZRKs across angiosperm diversity, we cloned wild-type *ZAR1* and *ZRK* genes from taro, stout camphor, and columbine. We also generated autoactive ZAR1 mutants by introducing a D to V mutation in the MHD motif following the approach we previously used for NbZAR1 ($NbZAR1^{D481V}$; Harant et al. 2022). In a series of experiments, we expressed the *ZAR1* and *ZRK* genes separately or in combination.

First, we examined whether wild-type and MHD mutants of ZAR1 are autoactive in *N. benthamiana*. We expressed wild-type and MHD mutant of taro ZAR1 ($CeZAR1^{WT}$, $CeZAR1^{D487V}$), stout camphor ZAR1 ($CmZAR1^{WT}$, $CmZAR1^{D488V}$), and columbine ZAR1 ($AcZAR1^{WT}$, $AcZAR1^{D489V}$) to determine whether wild-type and MHD mutant of these ZAR1 orthologs cause autoactive cell death in *N. benthamiana*. The 3 orthologs behaved differently in these assays. Whereas both $AcZAR1^{WT}$ and $AcZAR1^{D489V}$ induced autoactive cell death in *N. benthamiana* leaves, only the D to V mutant of CeZAR1 ($CeZAR1^{D487V}$) elicited cell death and neither 1 of $CmZAR1^{WT}$ and $CmZAR1^{D488V}$ caused a cell death response (Supplemental Fig. S7). As controls, we expressed wild-type and D to V mutant of Arabidopsis ZAR ($AtZAR1^{WT}$, $AtZAR1^{D489V}$) and *N. benthamiana* ZAR1 ($NbZAR1^{WT}$, $NbZAR1^{D481V}$). As reported previously (Baudin et al. 2019; Harant et al. 2022), $NbZAR1^{D481V}$ triggered autoactive cell death in *N. benthamiana* leaves but $AtZAR1^{D489V}$ did not (Supplemental Fig. S7).

Next, to determine whether wild-type ZRKs from taro, stout camphor, and columbine trigger autoactive cell death in *N. benthamiana*, we screened 19 ZRKs from taro (CeZRK1), stout camphor (CmZRK2, CmZRK3, CmZRK4, CmZRK5, CmZRK6, CmZRK7, CmZRK8, CmZRK9, CmZRK10, CmZRK11, CmZRK12, CmZRK13, and CmZRK15), and columbine (AcZRK1, AcZRK2, AcZRK3, AcZRK4, and AcZRK5; Pai et al. 2023). None of the tested ZRKs triggered macroscopic cell death response when expressed in *N. benthamiana* leaves (Supplemental Fig. S8; Pai et al. 2023). These results indicate that taro, stout camphor, and columbine ZRKs do not have autoactivity in *N. benthamiana*. This provides an opportunity to investigate functional connection between cospecific ZAR1 and ZRK orthologs by determining the effect of ZRK expression on ZAR1-mediated cell death response.

Stout camphor and columbine ZED1-related kinase (ZRK) proteins positively regulate the autoactive cell death of their cospecific ZAR1

To determine ZRK function in ZAR1-mediated cell death, we coexpressed D to V mutant versions of ZAR1 orthologs, $CeZAR1^{D487V}$, $CmZAR1^{D488V}$, and $AcZAR1^{D489V}$ with ZRK genes from each species. In these assays, CeZRK1 expression did not enhance cell death autoactivity of $CeZAR1^{D487V}$, whereas AcZRK1, AcZRK3, AcZRK4, and AcZRK5, but not AcZRK2, enhanced the cell death response of $AcZAR1^{D489V}$ (Fig. 6, A and B, and Supplemental Fig. S8). Coexpression of $CmZAR1^{D488V}$ together with CmZRK2, CmZRK6, CmZRK8, CmZRK9, CmZRK10, CmZRK11, or CmZRK13 caused macroscopic cell death in *N. benthamiana* leaves even though $CmZAR1^{D488V}$ itself did not trigger visible cell death (Fig. 6, C and D).

We further conducted side-by-side experiments coexpressing ZAR1 D to V mutants and ZRKs in comparison with single gene expression of either ZAR1 D to V mutants or ZRKs (Supplemental Fig. S9). This confirmed that 4 columbine ZRKs (AcZRK1, AcZRK3, AcZRK4, and AcZRK5) and 7 stout camphor ZRKs (CmZRK2, CmZRK6, CmZRK8, CmZRK9, CmZRK10, CmZRK11, and CmZRK13) positively regulate cell death activity of their cospecific ZAR1, although the ZRKs themselves did not show autoactivity in *N. benthamiana* (Supplemental Fig. S9). These results indicate that the ZAR1 orthologs of these species are functionally associated with ZRKs as previously shown for Arabidopsis ZAR1 and *N. benthamiana* ZAR1. We conclude that ZAR1 has been partnering with RLCKs for over 150 Mya of angiosperm evolution.

Considering that the interaction surfaces between ZAR1 and ZRKs are well conserved (Fig. 4A), we hypothesized that ZAR1 and ZRK proteins may be functionally interchangeable between different plant species. To test this, we coexpressed Arabidopsis RKS1/ZRK1 with D to V mutant of Arabidopsis ZAR1 (AtZAR1), AcZAR1, and CmZAR1 in the *N. benthamiana zar1-1* mutant line. As observed in the original ZAR1–ZRK experiments (Fig. 6), RKS1/ZRK1 positively regulated autoactive cell death by $CmZAR1^{D488V}$ (Supplemental Fig. S10). In the control experiment expressing $AtZAR1^{D489V}$ and RKS1/ZRK1, RKS1/ZRK1 conferred autoactivity to the AtZAR1 MHD mutant in the *N. benthamiana zar1-1* line (Supplemental Fig. S10). These experiments further confirm that the immune function of ZAR1 and ZRK family proteins is conserved across different flowering plant species.

Our observation that the CeZAR1 autoactive mutant triggered cell death regardless of CeZRK1 raised the possibility that CeZAR1 functions together with the endogenous JIM2 RLCK in *N. benthamiana*. To test this, we used a hairpin-silencing construct of *JIM2* (RNAi:JIM2) that mediates silencing of *JIM2* when transiently expressed in *N. benthamiana* leaves (Harant et al. 2022). Silencing of endogenous *JIM2* did not affect the cell death activity of $CeZAR1^{D487V}$, although it suppressed cell death triggered by $NbZAR1^{D481V}$ (Supplemental Fig. S11). This result indicates that unlike *N. benthamiana* ZAR1, taro ZAR1 triggers autoactive cell death independently of JIM2.

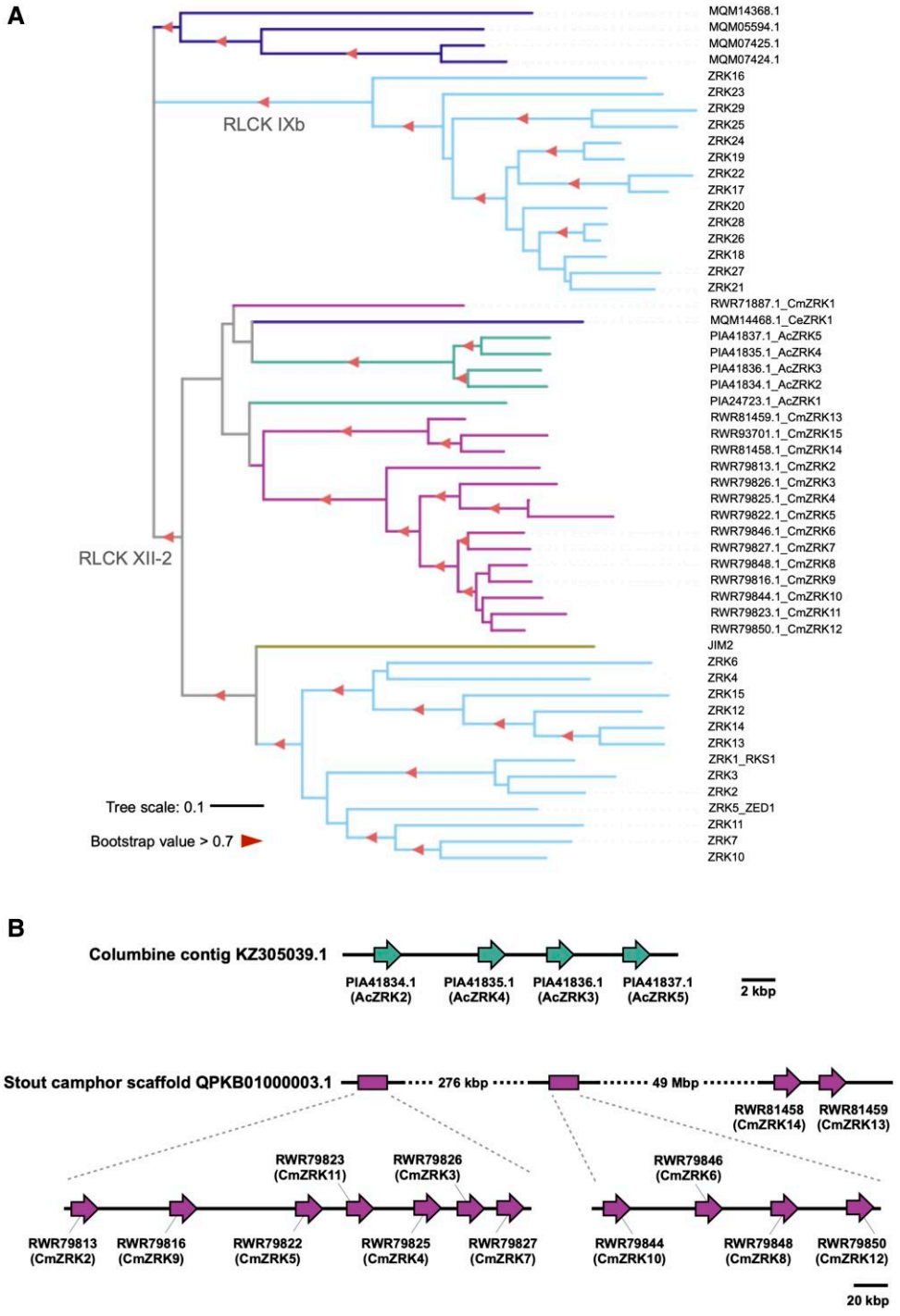


Figure 5. *ZRK* gene clusters occur in *A. coerulea* and *C. micranthum*. **A)** The phylogenetic tree was generated in MEGA7 by the neighbor-joining method using full-length amino acid sequences of 39 ZRK proteins. Red triangles indicate bootstrap support > 0.7. The scale bar indicates the evolutionary distance in amino acid substitution per site. **B)** Schematic representation of the *ZRK* gene clusters on an *A. coerulea* (columbine) contig and a *C. micranthum* (Stout camphor) scaffold.

Integration of a PLP3a thioredoxin-like domain at the C-termini of cassava and cotton ZAR1

As noted earlier, 9 ZAR1 orthologs carry an ID at their C-termini (Supplemental Data Set 1). These ZAR1-ID include 2 predicted proteins (XP_021604862.1 and XP_021604864.1) from *M. esculenta* (cassava) and 7 predicted proteins (KAB1998109.1, PPD92094.1, KAB2051569.1, TYG89033.1, TYI49934.1, TYJ04029.1, and KJB48375.1) from the cotton plant species *Gossypium barbadense*, *Gossypium darwinii*, *Gossypium mustelinum*, and *Gossypium raimondii* (Supplemental Data Set 1). The integrations follow an intact LRR domain, and the IDs vary in length from 108 to 266 amino acids (Fig. 7A). We

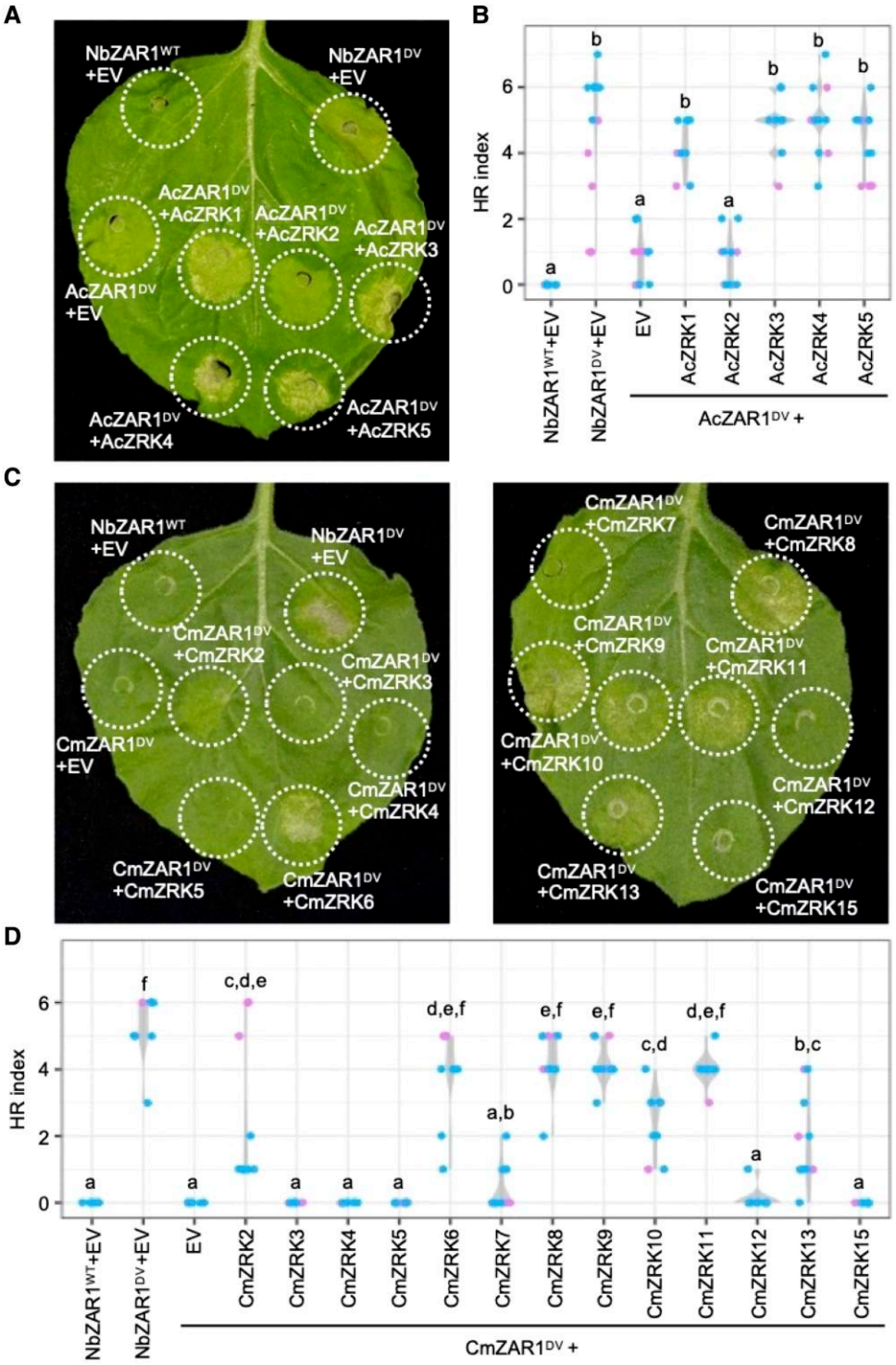


Figure 6. ZRK family proteins positively regulate *A. coerulea* AcZAR1 and *C. micranthum* CmZAR1 autoimmune cell death in *N. benthamiana*. **A, C)** Cell death observed in *N. benthamiana* after expression of ZAR1 mutants with or without wild-type ZRKs. *N. benthamiana* leaf panels expressing wild-type NbZAR1 ($NbZAR1^{WT}$), $NbZAR1^{D481V}$ ($ZAR1^{D481V}$), $AcZAR1^{D489V}$ ($AcZAR1^{DV}$), and $CmZAR1^{D488V}$ ($CmZAR1^{DV}$) with or without wild-type ZRKs, were photographed at 4 d after agroinfiltration. **B, D)** Violin plots show AcZAR1 and CmZAR1 cell death intensity scored as an HR index based on 12 and 9 replicates (different leaves from independent plants) in 2 independent experiments. Statistical differences among the samples were analyzed with Tukey's HSD test ($P < 0.01$).

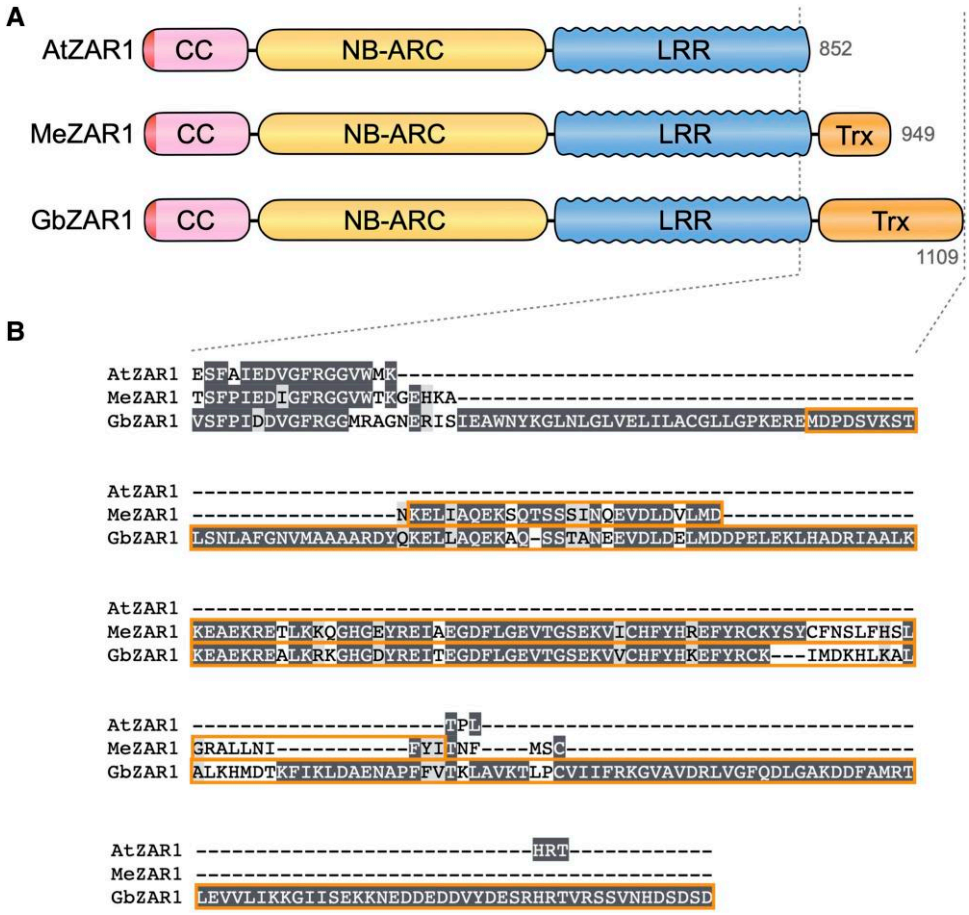


Figure 7. Cassava and cotton ZAR1-ID carry an additional Trx domain at the C terminus. **A)** Schematic representation of NLR domain architecture with C-terminal Trx domain. **B)** Description of Trx domain sequences on amino acid sequence alignment. Cassava XP_021604862.1 (MeZAR1) and cotton KAB1998109.1 (GbZAR1) were used for MAFFT version 7 alignment as representative ZAR1-ID. Arabidopsis ZAR1 (AtZAR1) was used as a control of ZAR1 without ID.

confirmed that the ZAR1-ID gene models of cassava XP_021604862.1 and XP_021604864.1 are correct based on RNA sequencing (RNA-seq) exon coverage in the NCBI database (database ID: LOC110609538). However, cassava *ZAR1-ID* XP_021604862.1 and XP_021604864.1 are isoforms encoded by transcripts from a single locus on chromosome LG2 (RefSeq sequence NC_035162.1) of the cassava RefSeq assembly (GCF_001659605.1) which also produces transcripts encoding isoforms lacking the C-terminal ID (XP_021604863.1, XP_021604865.1, XP_021604866.1, XP_021604867.1, and XP_021604868.1). Thus, cassava *ZAR1-ID* are probably splicing variants from a unique cassava *ZAR1* gene locus.

To determine whether the *ZAR1-ID* transcript is expressed in cassava, we analyzed public RNA-seq data from cassava samples in detail (BioSample IDs in NCBI database: SAMN02950671, SAMN02950673, SAMN02950674, SAMN02950672 SAMN02444910, SAMN02444915, AMN02444919, and SAMN05208186). We confirmed that RNA-seq reads detected from leaf and stem samples of 60,444 and MCOL1522 cassava cultivars span between the end of LRR and beginning of Trx domain regions (Supplemental Fig. S12). Notably, those reads are detected in the samples inoculated with the bacteria *Xanthomonas euvesicatoria* or *Xanthomonas axonopodis*, but not in their control samples (Supplemental Fig. S12). Furthermore, 3 reads spanning LRR and Trx regions are detected in RNA-seq data from a lateral bud of the TME204 cultivar (Supplemental Fig. S12). This suggests that ZAR1-ID is a splicing variant produced in cassava leaves and stems during *Xanthomonas* infection or in the specific tissue like the lateral bud.

To determine the phylogenetic relationship between ZAR1-ID and canonical ZAR1, we mapped the domain architectures of ZAR1 orthologs on the phylogenetic tree shown in Fig. 2 (Supplemental Fig. S13). Cassava and cotton *ZAR1-ID* occur in different branches of the ZAR1 rosid clade indicating that they may have evolved as independent integrations although alternative evolutionary scenarios such as a common origin followed by subsequent deletion of the ID or lineage sorting remain possible (Supplemental Fig. S13).

We annotated all the C-terminal extensions as thioredoxin-like using InterProScan (Trx, IPR036249; IPR013766; cd02989). The integrated Trx domain sequences are similar to Arabidopsis AT3G50960 (phosphoducin-like PLP3a; 34.8% to

90% similarity to integrated Trx domains), which is located immediately downstream of *ZAR1* in a tail-to-tail configuration in the Arabidopsis genome (Supplemental Fig. S14). We also noted additional genetic linkage between *ZAR1* and *Trx* genes in other rosid species, namely, field mustard, orange, cacao, grapevine, and apple, and in the asterid species coffee (Supplemental Data Set 5). We conclude that *ZAR1* is often genetically linked to a PLP3a-like Trx domain gene and that the ID in *ZAR1-ID* has probably originated from a genetically linked sequence.

The *ZAR1-SUB* clade emerged early in eudicot evolution from a single *ZAR1* duplication event

Phylogenetic analyses revealed *ZAR1-SUB* as a sister clade of the *ZAR1* ortholog clade (Figs. 1B and 8). The *ZAR1-SUB* clade comprises 129 genes from a total of 55 plant species (Supplemental Data Set 6). Twenty-one of the 55 plant species carry a single copy of *ZAR1-SUB* whereas 34 species have 2 or more copies (Supplemental Data Set 2). Of the 129 genes, 122 code for canonical CC-NLR proteins (692 to 1,038 amino acid length) with shared sequence similarities ranging from 36.5% to 99.9% (Supplemental Data Set 6).

Unlike *ZAR1*, *ZAR1-SUB* NLRs are restricted to eudicots (Supplemental Fig. S15 and Supplemental Data Set 6). Three out of 129 genes are from eudicot clade Ranunculales species outside the core eudicots, namely, columbine, *Macleaya cordata* (plume poppy) and *Papaver somniferum* (opium poppy). The remaining ZAR1-SUB are spread across rosid and asterid species. We found that 11 species have *ZAR1-SUB* genes but lack a *ZAR1* ortholog (Supplemental Data Set 2). These 11 species include 2 of the early diverging eudicots plume poppy and opium poppy, and the Brassicales *Carica papaya* (papaya). Interestingly, papaya is the only Brassicales species carrying a *ZAR1-SUB* gene, whereas the 16 other Brassicales species have *ZAR1* but lack *ZAR1-SUB* genes (Supplemental Data Set 2). In total, we didn't detect *ZAR1-SUB* genes in 44 species that have *ZAR1* orthologs and these 44 species include the monocot taro, the magnoliid stout camphor, and 42 eudicots, such as Arabidopsis, sugar beet, and *N. benthamiana* (Supplemental Data Set 2).

In summary, given the taxonomic distribution of the *ZAR1-SUB* clade genes, we propose that *ZAR1-SUB* has emerged from a single duplication event of *ZAR1* prior to the split between Ranunculales and other eudicot lineages about ∼120 to 130 Mya based on the species divergence time estimate of Chaw et al. (2019).

ZAR1-SUB paralogs have significantly diverged from *ZAR1*

We investigated the sequence patterns of ZAR1-SUB proteins and compared them to the sequence features of canonical ZAR1 proteins that we identified earlier (Fig. 3A). MEME analyses revealed several conserved sequence motifs (Supplemental Table S2). Especially, the MEME motifs in the ZAR1-SUB NB-ARC domain were similar to ZAR1 ortholog motifs (Supplemental Table S3). These include P-loop and MHD motifs, which are broadly conserved in NB-ARC of 97% and 100% of the ZAR1-SUB NLRs, respectively (Fig. 9A). MEME also revealed sequence motifs in the ZAR1-SUB LRR domain that partially overlaps in position with the conserved ZAR1–RLCK interfaces (Fig. 9A and Supplemental Fig. S16). However, the ZAR1-SUB MEME motifs in the LRR domain were variable at the ZAR1–RLCK interface positions compared to ZAR1 and the motif sequences were markedly different between ZAR1-SUB and ZAR1 proteins (Figs. 3A and 9A).

Remarkably, unlike ZAR1 orthologs, MEME did not predict the conserved sequence pattern from a region corresponding to the MADA motif, indicating that these sequences have diverged across ZAR1-SUB proteins (Fig. 9A). We confirmed the low frequency of MADA motifs in ZAR1-SUB proteins using HMMER searches with only ∼30% (38 out of 129) of the tested proteins having a MADA-like sequence (Supplemental Data Set 6; Fig. 8). Moreover, conserved sequence patterns were not predicted for the NBD–NBD interface of the ZAR1 resistosome (Fig. 9A and Supplemental Fig. S16).

We generated a diversity barcode for ZAR1-SUB proteins using ConSurf as we did earlier with ZAR1 orthologs (Fig. 9B). This revealed that there are several conserved sequence blocks in each of the CC, NB-ARC, and LRR domains, such as the regions corresponding to P-loop, MHD motif, and the equivalent of the ZAR1–RLCK interfaces.

Next, we mapped the ConSurf conservation scores onto a homology model of a representative ZAR1-SUB protein (XP_004243429.1 from tomato) built based on the Arabidopsis ZAR1 cryo-EM structures (Supplemental Fig. S17). As highlighted in Supplemental Fig. S17B and C, conserved residues, such as the MHD motif region in the WHD, are located inside of the monomer and resistosome structures. Interestingly, although the prior MEME prediction analyses revealed conserved motifs in positions matching the ZAR1–RLCK interfaces in the LRR domain, the ZAR1-SUB structure homology models displayed variable surfaces in this region (Supplemental Fig. S17A). This indicates that the variable residues within these sequence motifs are predicted to be on the outer surfaces of the LRR domain and may reflect interaction with different ligands.

Taken together, these results suggest that unlike ZAR1 orthologs, the ZAR1-SUB paralogs have divergent molecular patterns for regions known to be involved in effector recognition, resistosome formation, and activation of hypersensitive cell death.

Eleven tandemly duplicated *ZAR1-CIN* genes occur in a 500-kb cluster in the *C. micranthum* (stout camphor) genome

The *ZAR1-CIN* clade, identified by phylogenetic analyses as a sister clade to *ZAR1* and *ZAR1-SUB*, consists of 11 genes from the magnoliid species stout camphor (Figs. 1B and 8 and Supplemental Data Set 7). Eight of the 11 *ZAR1-CIN* genes code for canonical CC-NLR proteins with 63.8% to 98.9% sequence similarities to each other, whereas the remaining 3 genes code for truncated

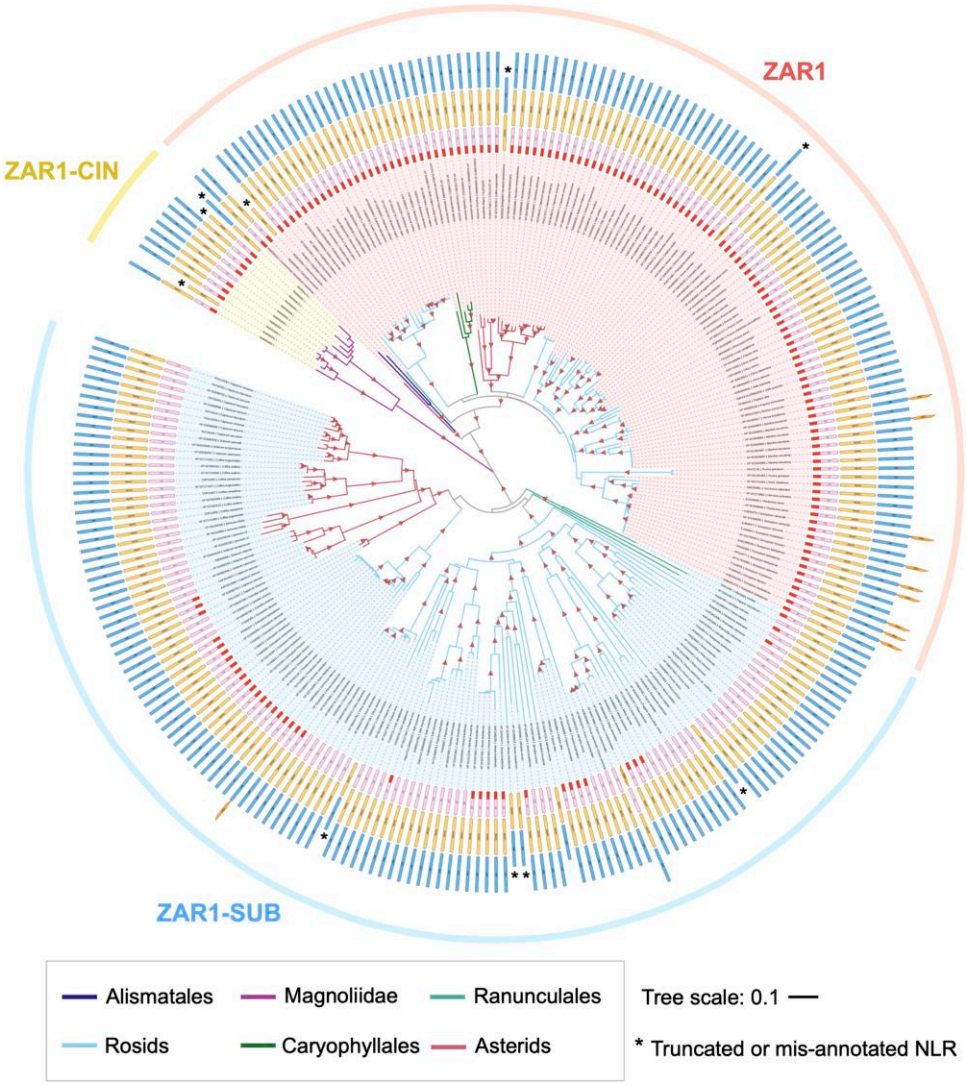


Figure 8. ZAR1-SUB has emerged early in eudicots and diverged at MADA motif sequence. The phylogenetic tree was generated in MEGA7 by the neighbor-joining method using full-length amino acid sequences of 120 ZAR1, 129 ZAR1-SUB, and 11 ZAR1-CIN identified in Fig. 1. Red triangles indicate bootstrap support > 0.7. The scale bar indicates the evolutionary distance in amino acid substitution per site. NLR domain architectures are illustrated outside of the leaf labels: MADA is red, CC is pink, NB-ARC is yellow, LRR is blue, and other domain is orange. Black asterisks on domain schemes describe truncated NLRs or potentially misannotated NLR.

NLR proteins. Interestingly, all *ZAR1-CIN* genes occur in a ~500-kb cluster on scaffold QPKB01000005.1 of the stout camphor genome assembly (GenBank assembly accession GCA_003546025.1) (Supplemental Fig. S18, A and B). This scaffold also contains the stout camphor *ZAR1* ortholog (*CmZAR1*, RWR84015), which is located 48 Mb from the *ZAR1-CIN* cluster (Supplemental Fig. S18, A and B). Based on the observed phylogeny and gene clustering, we suggest that the *ZAR1-CIN* cluster emerged from segmental duplication and expansion of the ancestral *ZAR1* gene after stout camphor split from the other examined *ZAR1*-containing species.

We examined the expression of the eleven *CmZAR1* and *ZAR1-CIN* genes in 7 tissues of *C. micranthum* based on the data of Chaw et al. (Chaw et al. 2019). The *CmZAR1* gene is relatively highly expressed in 7 different tissues of the stout camphor tree (Supplemental Fig. S18C). In contrast, only 5 of the eleven *ZAR1-CIN* genes displayed detectable expression levels. Of these, 2 *ZAR1-CIN* genes (RWR85656 and RWR85657) had different expression patterns across the tissues. Whereas RWR85657 had the highest expression level in flowers, RWR85656 displayed the highest expression levels in stem and old leaf tissues (Supplemental Fig. S18C). The implications of these observations remain unclear but may reflect different degrees of tissue specialization of the *ZAR1-CIN* genes.

Tandemly duplicated *ZAR1-CIN* genes display variable ligand-binding interfaces on the LRR domain

We performed MEME and ConSurf analyses of the 8 intact ZAR1-CIN proteins as described above for ZAR1 and

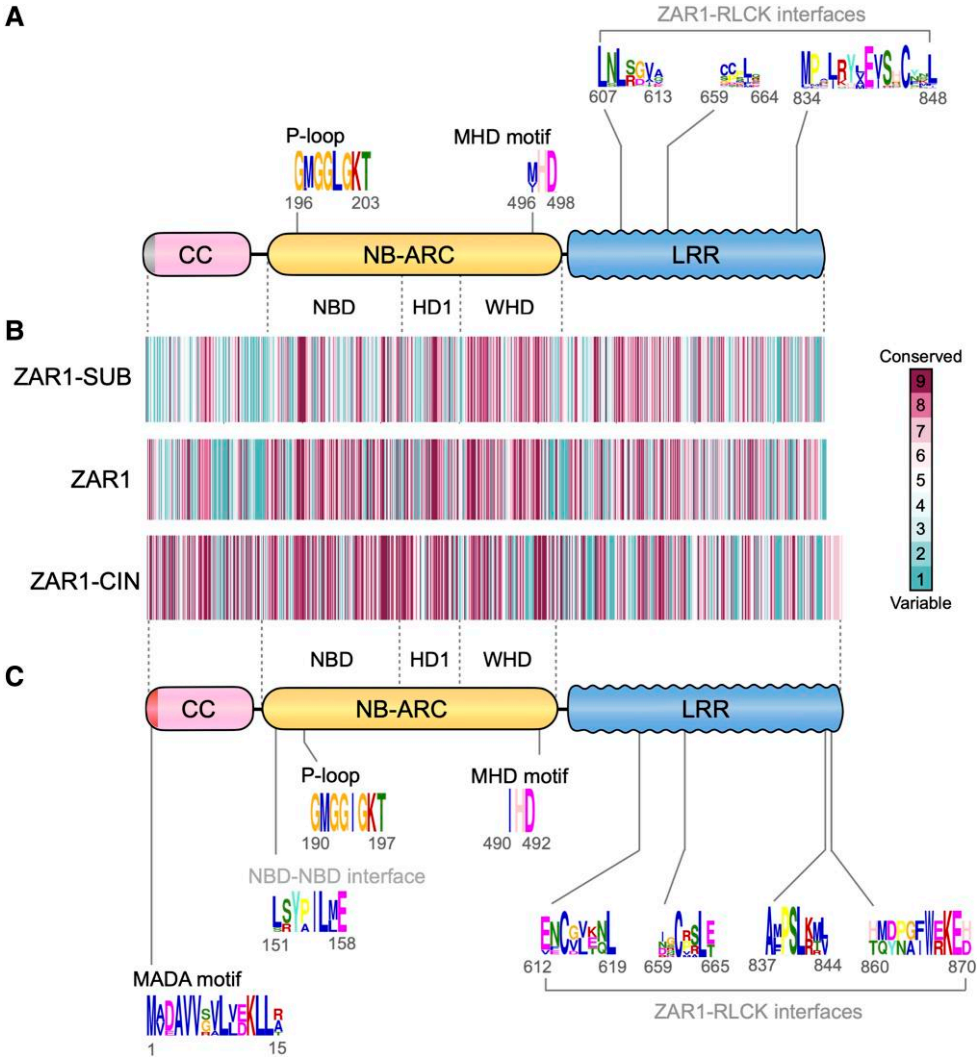


Figure 9. Conserved sequence distributions in ZAR1-SUB and ZAR1-CIN. **A)** Schematic representation of the ZAR1-SUB protein highlighting the position of the representative conserved sequence patterns across ZAR1-SUB. Representative consensus sequence patterns identified by MEME are described on the scheme. Raw MEME motifs are listed in Supplemental Tables S2 and S3. **B)** Conservation and variation of each amino acid among ZAR1-SUB and ZAR1-CIN. Amino acid alignment of 129 ZAR1-SUB or 8 ZAR1-CIN was used for conservation score calculation via the ConSurf server (https://consurf.tau.ac.il). The conservation scores are mapped onto each amino acid position in queries XP_004243429.1 (ZAR1-SUB) and RWR85656.1 (ZAR1-CIN), respectively. **C)** Schematic representation of the ZAR1-CIN protein highlighting the position of the representative conserved sequence patterns across 8 ZAR1-CIN. Raw MEME motifs are listed in Supplemental Tables S4 and S5.

ZAR1-SUB. The ConSurf barcode revealed that although ZAR1-CIN proteins are overall conserved, their WHD region and LRR domain include some clearly variable blocks (Fig. 9B). MEME analyses of ZAR1-CIN sequences revealed that like ZAR1 orthologs, the MADA, P-loop, and MHD motifs match highly conserved blocks of the ZAR1-CIN ConSurf barcode (Fig. 9, B and C, and Supplemental Tables S4 and S5). Consistently, 87.5% (7 out of 8) of the ZAR1-CIN proteins were predicted to have a MADA-type N-terminal sequence based on MADA-HMM analyses (Supplemental Data Set 7 and Fig. 8).

MEME picked up additional sequence motifs in ZAR1-CIN proteins that overlap in position with the NBD–NBD and ZAR1–RLCK interfaces (Fig. 9C and Supplemental Fig. S19). However, the sequence consensus at the NBD–NBD and ZAR1–RLCK interfaces indicated these motifs are more variable among ZAR1-CIN proteins relative to ZAR1 orthologs and the motif sequences at both interfaces were markedly different from the matching region in ZAR1 (Figs. 3A and 9C).

We also mapped the ConSurf conservation scores onto a homology model of a representative ZAR1-CIN protein (RWR85656.1) built based on the Arabidopsis ZAR1 cryo-EM structures (Supplemental Fig. S17). This model revealed several conserved surfaces, such as on the α1 helix in the CC domain and the WHD of the NB-ARC domain (Supplemental Fig. S17, B and C). In contrast, the ZAR1-CIN structure homology models displayed highly varied surfaces especially in the LRR region matching the

RLCK binding interfaces of ZAR1 (Supplemental Fig. S17A). This sequence diversification on the LRR surface suggests that the ZAR1-CIN paralogs may have different host partner proteins and/or effector recognition specificities compared to ZAR1.

Discussion

This study of ZAR1 macroevolution originated from phylogenomic analyses we initiated during the UK COVID-19 lockdown of March 2020. We performed iterated comparative sequence similarity searches of plant genomes using the CC-NLR immune receptor ZAR1 as a query, and subsequent phylogenetic evaluation of the recovered ZAR1-like sequences. This revealed that *ZAR1* is an ancient gene with 120 orthologs recovered from 88 species including monocot, magnoliid, and eudicot plants. This atypical conservation of *ZAR1* in these species provides the view that *ZAR1* originated early in angiosperms during the Jurassic geologic period ~220 to 150 Mya (Fig. 10A). This evolutionary model of ZAR1 is consistent with a recent study by Gong et al. (2022) that was published 1.5 yrs after we posted a preprint of the present paper (Adachi et al. 2020). The ortholog series enabled us to determine that resistosome sequences that are known to be functionally important and have remained highly conserved throughout the long evolutionary history of ZAR1. In addition, we experimentally validated the model that ZAR1 has partnered with RLCKs for over 150 Mya through functional reconstruction of ZAR1–RLCK pairs from distantly related plant species (Fig. 10B). The main unexpected feature among ZAR1 orthologs is the acquisition of a C-terminal thioredoxin-like domain in cassava and cotton species (Fig. 7). Our phylogenetic analyses also indicated that *ZAR1* duplicated twice throughout its evolution (Fig. 10A). In the eudicots, *ZAR1* spawned a large paralog family, *ZAR1-SUB*, which greatly diversified and often lost the typical sequence features of *ZAR1*. A second paralog, *ZAR1-CIN*, is restricted to a tandemly repeated 11-gene cluster in stout camphor. Overall, our findings map patterns of functional conservation, expansion, and diversification onto the evolutionary history of *ZAR1* and its paralogs.

ZAR1 most likely emerged prior to the split between monocots, Magnoliids, and eudicots, which corresponds to ~220 to 150 Mya based on the dating analyses of Chaw et al. (2019). The origin of the angiosperms remains hotly debated with uncertainties surrounding some of the fossil record coupled with molecular clock analyses that would benefit from additional genome sequences of undersampled taxa (Coiro et al. 2019). Fu et al. (2018) and Cui et al. (2022) provided credence to an earlier emergence of angiosperms with the discovery of the fossil flowers *Nanjinganthus dendrostyla* and *Florigerminis jurassica*, respectively. These findings place the emergence of flowering plants at the Jurassic. It is tempting to speculate that *ZAR1* emerged among these early flowering plants during the period when dinosaurs dominated planet Earth.

NLRs are notorious for their rapid and dynamic evolutionary patterns even at the intraspecific level. In sharp contrast, *ZAR1* is an atypical core NLR gene conserved in a wide range of angiosperm species (Fig. 2). Nevertheless, Arabidopsis ZAR1 can recognize diverse bacterial pathogen effectors, including 5 different effector families distributed among nearly half of a collection of ~500 *Pseudomonas syringae* strains (Laflamme et al. 2020) and an effector AvrAC from *Xanthomonas campestris* (Wang et al., 2015). How did ZAR1 remain conserved throughout its evolutionary history while managing to detect a diversity of effectors? The answer to the riddle lies in the fact that ZAR1 effector recognition occurs via its partner RLCKs. ZRKs of the RLCK XII-2 subfamily rest in complex with inactive ZAR1 proteins and bait effectors by binding them directly or by recruiting other effector-binding RLCKs, such as family VII PBS1-like protein 2 (PBL2) (Lewis et al. 2013; Wang et al., 2015). These ZAR1-associated RLCKs are highly diversified not only in Arabidopsis (Lewis et al. 2013) but also in stout camphor and columbine, where RLCK XII-2 members occur in expanded ZRK gene clusters (Fig. 5). In the Arabidopsis ZRK cluster, RKS1/ZRK1 is required for recognition of the *X. campestris* effector AvrAC (Wang et al., 2015) and ZRK3 and ZED1/ZRK5 are required for recognition of *P. syringae* effectors HopF2a and HopZ1a, respectively (Lewis et al. 2013; Seto et al. 2017). Therefore, as in the model discussed by Schultink et al. (2019) and Gong et al. (2022), ZRKs appear to have evolved as pathogen "sensors" whereas ZAR1 acts as a conserved signal executor to activate immune response.

The MEME and ConSurf analyses are consistent with the model of ZAR1/RLCK evolution described above. *ZAR1* is not just exceptionally conserved across angiosperms, but it has also preserved sequence patterns that are key to resistosome-mediated immunity (Figs. 3 and 4). Within the LRR domain, ZAR1 orthologs display highly conserved surfaces for RLCK binding (Fig. 4). We conclude that ZAR1 has been guarding host kinases throughout its evolution ever since the Jurassic period. These findings strikingly contrast with observations recently made by Prigozhin and Krasileva (2020) on highly variable Arabidopsis NLRs (hvNLRs), which tend to have diverse LRR sequences. For instance, the CC-NLR RPP13 displays variable LRR surfaces across 62 Arabidopsis accessions, presumably because these regions are effector recognition interfaces that are caught in arms race coevolution with the oomycete pathogen *Hyaloperonospora arabidopsidis* (Prigozhin and Krasileva 2020). The emerging view is that the mode of pathogen detection (direct vs indirect recognition) drives an NLR evolutionary trajectory by accelerating sequence diversification at the effector binding site or by maintaining the binding interface with the partner guardee/decoy proteins (Prigozhin and Krasileva 2020).

Our functional validation of ZAR1 and ZRKs from distantly related plant species supported the model that ZRKs function together with ZAR1 to trigger immune response *in planta* (Fig. 6). Eleven of the 19 tested ZRKs either required or enhanced the autoactivity of their cospecific ZAR1 in *N. benthamiana*. The

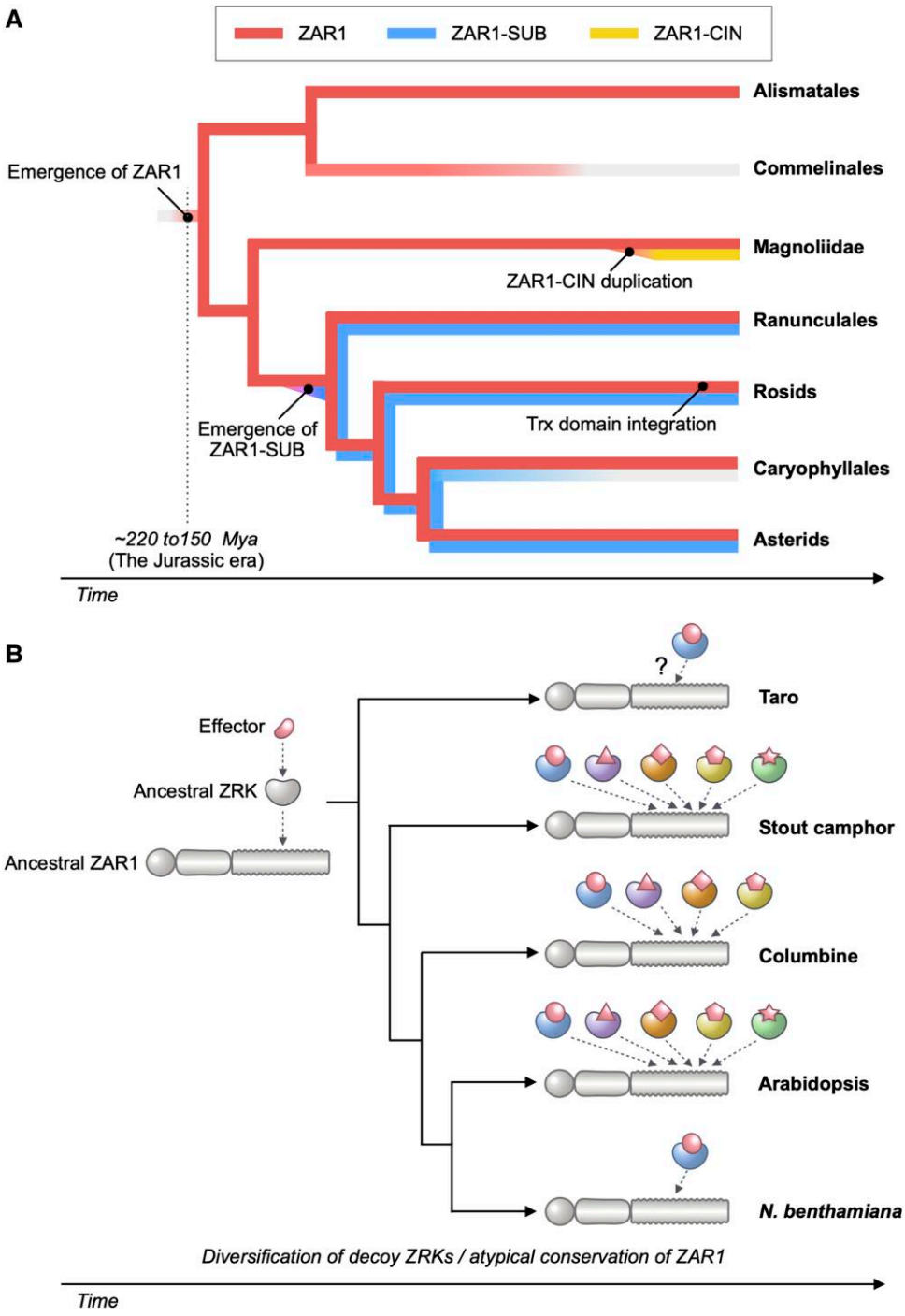


Figure 10. Coevolution of *ZAR1* and *ZRK* genes in angiosperms. **A)** We propose that the ancestral *ZAR1* gene has emerged ~220 to 150 million yrs ago (Mya) before monocot and eudicot lineages split. ZAR1 is a widely conserved CC-NLR in angiosperms, but it is likely that *ZAR1* was lost in the monocot lineage, Commelinales. A sister clade paralog *ZAR1-SUB* has emerged early in the eudicot lineages and may have been lost in Caryophyllales. Another sister clade paralog *ZAR1-CIN* was duplicated from the *ZAR1* gene and expanded in the Magnoliidae *C. micranthum*. Trx domain integration to C terminus of ZAR1 has independently occurred in few rosid lineages. **B)** *ZAR1* has coevolved with partner *ZRK* gene for pathogen effector recognition since the Jurassic era. During the coevolution, ZRKs diversified to catch up with fast-evolving effectors.

remaining 8 tested ZRKs, CeZRK1, AcZRK2, CmZRK3, CmZRK4, CmZRK5, CmZRK7, CmZRK12, and CmZRK15 did not alter cell death activity of ZAR1. Notably, CmZRK4, CmZRK5, and CmZRK12 have N-terminal truncation or mutations at the ZAR1 interaction sites identified from the Arabidopsis ZAR1–ZRK studies (Supplemental File 2). Therefore, some of the ZRK members may have lost their association with ZAR1 through deletion or mutations.

Taro and columbine ZAR1 could trigger autoactivated cell death without their partner RLCKs in *N. benthamiana*, whereas stout camphor and *N. benthamiana* ZAR1 proteins require ZRKs to trigger the cell death response (Fig. 6 and Supplemental Fig. S7) (Harant et al. 2022). In the case of taro and columbine, ZRKs may trigger conformational changes of ZAR1 after recognition of cognate pathogen effectors. In this scenario, autoactive ZAR1 could form a resistosome without ZRK proteins, thereby triggering the observed cell death response. In the future, further comparative biochemical studies would further inform our understanding of how ZAR1–ZRK interactions have evolved and contributed to ZAR1 resistosome formation across angiosperms.

ZAR1 orthologs display a patchy distribution across angiosperms (Supplemental Data Set 1). Given the low number of non-eudicot species with *ZAR1*, it is challenging to develop a conclusive evolutionary model. Nonetheless, the most parsimonious explanation is that *ZAR1* was lost in the monocot Commelinales lineage (Fig. 10A and Supplemental Data Set 1). *ZAR1* is also missing in some eudicot lineages, notably Fabales, Cucurbitales, Apiales, and Asterales (Supplemental Data Set 1). Cucurbitaceae (Cucurbitales) species are known to have reduced repertoires of NLR genes possibly due to low levels of gene duplications and frequent deletions (Lin et al. 2013). *ZAR1* may have been lost in this and other plant lineages as part of an overall shrinkage of their NLRomes or as consequence of selection against autoimmune phenotypes triggered by NLR misregulation (Karasov et al. 2017; Adachi et al. 2019a). Notably, plant linages that don't have a *ZAR1* ortholog also lack *ZRK* family genes, suggesting that ZAR1 and ZRK coevolved to function in resistosome-mediated immunity across angiosperms (Gong et al. 2022).

We unexpectedly discovered that some *ZAR1* orthologs from cassava and cotton species carry a C-terminal thioredoxin-like domain (*ZAR1-ID* in Fig. 7). Although Gong et al. (2022) suggested *ZAR-ID* are annotation errors, we confirmed that at least cassava *ZAR1-Trx* is expressed as a splicing variant in leaf and stem inoculated with *Xanthomonas* bacteria or in lateral bud (Supplemental Fig. S12). What is the function of these IDs? The occurrence of unconventional domains in NLRs is relatively frequent and ranges from 5% to 10% of all NLRs. In several cases, IDs have emerged from pathogen effector targets and became decoys that mediate detection of the effectors (Kourelis and van der Hoorn 2018). Whether or not the integrated Trx domain of ZAR1-ID functions to bait effectors will need to be investigated. Since ZAR1-ID proteins still carry intact RLCK binding interfaces (Supplemental File 3), they may have evolved dual or multiple recognition specificities via RLCKs and the Trx domain. In addition, all ZAR1-ID proteins have an intact N-terminal MADA motif (Supplemental Fig. S13), suggesting that they probably can execute the hypersensitive cell death through their N-terminal CC domains even though they carry a C-terminal domain extension (Adachi et al. 2019b). In the future, it would be intriguing to understand how the ZAR1-ID splicing variant is produced and how the ZAR-ID function comparing to the ZAR1 resistosome model.

Our sequence analyses of *ZAR1-ID* indicate that the integrated Trx domain originates from the *PLP3* phosphoducin gene, which is immediately downstream of *ZAR1* in the Arabidopsis genome and adjacent to *ZAR1* in several other eudicot species (Supplemental Fig. S14). Whether or not PLP3 plays a role in ZAR1 function and the degree to which close genetic linkage facilitated domain fusion between these 2 genes are provocative questions for future studies.

ZAR1 spawned 2 classes of paralogs through 2 independent duplication events. The *ZAR1-SUB* paralog clade emerged early in the eudicot lineage—most likely tens of millions of years after the emergence of *ZAR1*—and has diversified into at least 129 genes in 55 species (Fig. 10A). ZAR1-SUB proteins are distinctly more diverse in sequence than ZAR1 orthologs and generally lack key sequence features of ZAR1, like the MADA motif and the NBD–NBD oligomerization interface (Fig. 9) (Adachi et al. 2019b; Wang et al., 2019b; Hu et al. 2020). This pattern is consistent with the "use-it-or-lose-it" evolutionary model, in which NLRs that specialize for pathogen detection lose some of the molecular features of their multifunctional ancestors (Adachi et al. 2019b). Therefore, we predict that many ZAR1-SUB proteins evolved into specialized sensor NLRs that require NLR helper mates for executing the hypersensitive response. It is possible that ZAR1-SUB helper mate is ZAR1 itself and that these NLRs evolved into a phylogenetically linked network of sensors and helpers similar to the NRC network of asterid plants (Wu et al. 2017). However, 11 species have a *ZAR1-SUB* gene but lack a canonical *ZAR1* (Supplemental Data Set 2), indicating that these ZAR1-SUB NLRs may have evolved to depend on other classes of NLR helpers.

How would ZAR1-SUB sense pathogens? Given that the LRR domains of most ZAR1-SUB proteins markedly diverged from the RLCK binding interfaces of ZAR1, it is unlikely that all of ZAR1-SUB members bind RLCKs in a ZAR1-type manner (Supplemental Fig. S17). This leads us to draw the hypothesis that ZAR1-SUB proteins have diversified to recognize other ligands than RLCKs. Indeed, Gong et al. (2022) showed that only *Populus trichocarpa* and *Prunus persica* ZAR1-SUB (PtZAR1-SUB and PpZAR1-SUB) out of 6 tested ZAR1-SUB members interacted with ZRK proteins in coimmunoprecipitation experiments. Both PtZAR1-SUB and PpZAR1-SUB did not form ZAR1-like oligomer complexes with RKS1 and did not cause cell death responses. Therefore, ZAR1-SUB may require other components to be activated or execute immune responses. In the future, functional investigations of ZAR1-SUB proteins could provide insights into how multifunctional NLRs, such as ZAR1, evolve into functionally specialized NLRs.

The *ZAR1-CIN* clade consists of 11 clustered paralogs that are unique to the magnoliid species stout camphor as revealed from the genome sequence of the Taiwanese small-flowered camphor tree (also known as *C. kanehirae*,

Chinese name niu zhang 牛樟) (Chaw et al. 2019). This cluster probably expanded from *ZAR1*, which is ~48 Mbp on the same genome sequence scaffold (Supplemental Fig. S18). The relatively rapid expansion of *ZAR1-CIN* into a tandemly duplicated gene cluster is more in line with the classical model of NLR evolution compared to *ZAR1* maintenance as a genetic singleton over tens of millions of years (Michelmore and Meyers 1998). ZAR1-CIN proteins may have neofunctionalized after duplication, acquiring new recognition specificities as consequence of coevolution with host partner proteins and/or pathogen effectors. Consistent with this view, *ZAR1-CIN* exhibit different patterns of gene expression across tissues (Supplemental Fig. S18). Moreover, ZAR1-CIN proteins display distinct surfaces at the ZAR1–RLCK binding interfaces and may bind to other ligands than RLCKs as we hypothesized above for ZAR1-SUB (Supplemental Fig. S16). ZAR1-CIN could be viewed as intraspecific highly variable NLRs (hvNLR) per the nomenclature of Prigozhin and Krasileva (2020).

Unlike ZAR1-SUB, ZAR1-CIN have retained the N-terminal MADA sequence (Fig. 9 and Supplemental Fig. S17). We propose that ZAR1-CIN are able to execute the hypersensitive cell death on their own similar to ZAR1. However, ZAR1-CIN display divergent sequence patterns at NBD–NBD oligomerization interfaces compared to ZAR1 (Fig. 9C and Supplemental Fig. S19). Therefore, ZAR1-CIN may form resistosome-type complexes that are independent of ZAR1. One intriguing hypothesis is that ZAR1-CIN may associate with each other to form heterocomplexes of varying complexity and functionality operating as an NLR receptor network. In any case, the clear-cut evolutionary trajectory from *ZAR1* to the *ZAR1-CIN* paralog cluster provides a robust evolutionary framework to study functional transitions and diversifications in this CC-NLR lineage.

In summary, our phylogenomics analyses raise several intriguing questions about ZAR1 evolution. The primary conclusion we draw is that ZAR1 is an ancient CC-NLR that has been a partner with RLCKs ever since the Jurassic period. We propose that throughout at least 150 million yrs, ZAR1 has maintained its molecular features for sensing pathogens via RLCKs and activating hypersensitive cell death. Further comparative analyses, combining molecular evolution and structural biology, of plant resistosomes and between resistosomes and the apoptosomes and inflammasome of animal NLR systems (Wang and Chai 2020) will yield experimentally testable hypotheses for NLR research.

Materials and methods

ZAR1 and ZRK sequence retrieval

We performed BLAST (Altschul et al. 1990) using previously identified ZAR1 and ZRK protein sequences as queries (Lewis et al. 2013; Baudin et al. 2017; Schultink et al. 2019; Harant et al. 2022) to search ZAR1 and ZRK-like sequences in NCBI nr or nr/nt database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and Phytozome12.1 (https://phytozome.jgi.doe.gov/pz/portal.html#!search? show=BLAST). In the BLAST search, we used cut-offs, percent identity ≥ 30%, and query coverage ≥ 80%. The BLAST pipeline was circulated by using the obtained sequences as new queries to search for *ZAR1* and *ZRK*-like genes over the angiosperm species. We also performed the BLAST pipeline against a plant NLR dataset annotated by the NLR-parser (Steuernagel et al. 2015) from 38 plant reference genome databases (Supplemental Data Set 8).

Phylogenetic analyses

For the phylogenetic analysis, we aligned NLR and ZRK amino acid sequences (Supplemental Files 1, 4 to 8) using MAFFT v.7 (Katoh and Standley 2013) and manually deleted the gaps in the alignments in MEGA7 (Kumar et al. 2016). Full-length or NB-ARC domain sequences of the aligned NLR datasets were used for generating phylogenetic trees. To generate ZRK phylogenetic trees, we used full-length or kinase domain sequences of the aligned ZRK datasets. The neighbor-joining tree was made using MEGA7 with JTT model and bootstrap values based on 100 iterations. All phylogenetic tree files are in Supplemental Files 9 to 13.

Patristic distance analyses

To calculate the phylogenetic (patristic) distance, we used Python script based on DendroPy (Sukumaran and Holder 2010). We calculated patristic distances from each CC-NLR to the other CC-NLRs on the phylogenetic tree and extracted the distance between CC-NLRs of Arabidopsis (*A. thaliana*) or *N. benthamiana* to the closest NLR from the other plant species. The script used for the patristic distance calculation is available from GitHub (https://github.com/slt666666/Phylogenetic_distance_plot2).

Gene colinearity analyses

To investigate genetic colinearity at *ZAR1* loci, we extracted the 3 genes upstream and downstream of *ZAR1* using GFF files derived from reference genome databases (Supplemental Data Set 8). To identify conserved gene blocks, we used gene annotation from NCBI Protein database and confirmed protein domain information based on InterProScan (Jones et al, 2014).

Sequence conservation analyses

Full-length NLR sequences of each subfamily ZAR1, ZAR1-SUB, or ZAR1-CIN were subjected to motif searches using the MEME (Bailey and Elkan 1994) with parameters "0 or 1 occurrence per sequence, top twenty motifs," to detect consensus motifs conserved in ≥90% of the input sequences. The output data are summarized in Supplemental Tables S1, S2, and S4.

To predict the MADA motif from ZAR1, ZAR1-SUB, and ZAR1-CIN datasets, we used the MADA-HMM previously developed (Adachi et al. 2019b), with the hmmsearch program (hmmsearch –max -o <outputfile> <hmmfile>) implemented in HMMER v2.3.2 (Eddy 1998). We termed sequences over the HMMER cut-off score of 10.0 as the MADA motif and sequences having the score 0-to-10.0 as the MADA-like motif.

To analyze sequence conservation and variation in ZAR1, ZAR1-SUB, and ZAR1-CIN proteins, aligned full-length NLR sequences (Supplemental Files 3, 14, and 15) were used for ConSurf (Ashkenazy et al. 2016). Arabidopsis ZAR1 (NP_190664.1), a tomato ZAR1-SUB (XP_004243429.1), or a Stout camphor ZAR1-CIN (RWR85656.1) was used as a query for each analysis of ZAR1, ZAR1-SUB, or ZAR1-CIN, respectively. The output datasets are in Supplemental Data Sets 9 to 11.

Protein structure analyses

The atomic coordinate of ZAR1 (protein data bank accession codes; 6J5T) was downloaded from protein data bank for illustration in ccp4mg. We used the cryo-EM structures of ZAR1 as templates to generate homology models of ZAR1-SUB and ZAR1-CIN. Amino acid sequences of a tomato ZAR1-SUB (XP_004243429.1) and a stout camphor ZAR1-CIN (RWR85656.1) were submitted to Protein Homology Recognition Engine V2.0 (Phyre2) for modeling (Kelley et al., 2015). The coordinates of the ZAR1 structure (6J5T) were retrieved from the Protein Data Bank and assigned as the modeling template by using Phyre2 Expert Mode. The resulting models of ZAR1-SUB and ZAR1-CIN and the ZAR1 structures (6J5T) were illustrated with the ConSurf conservation scores in PyMol.

Plant growth conditions

Wild-type *N. benthamiana* and *zar1-1* mutant plants (Schultink et al. 2019) were grown in a controlled growth chamber with temperature 22 to 25°C, humidity 45% to 65%, and 16/8 hr light/dark cycle. Fluorescent light bulbs (Slyvania Gro—Lux F58W/Gro—T8 and Phillips master TL-D 58W84D) were used, and the light intensity was about 200 μM/ms^2.

Plasmid construction

The Golden Gate Modular Cloning (MoClo) kit (Weber et al. 2011) and the MoClo plant parts kit (Engler et al. 2014) were used for cloning, and all vectors are from this kit unless specified otherwise. ZAR1 and RLCK homologs identified in the taro (*C. esculenta*; Assembly: ASM944546v1), columbine (*A. coerulea*; Assembly: Aquilegia_coerulea_v1; Filiault et al. 2018), and stout camphor tree (*C. kanehirae*; Assembly: ASBRC_Ckan_1.0; Chaw et al. 2019) genomes were codon optimized for *N. benthamiana* using the ThermoFisher GeneOptimizer tool and synthesized by GENEWIZ as Golden Gate Level 0 modules into pICH41155. Genes were subcloned into the binary vector pICH86988 (Weber et al. 2011) and transformed into *A. tumefaciens* strain GV3101 pMP90. Cloning design and sequence analysis were done using Geneious Prime (v2022.0.1; https://www.geneious.com). Plasmid construction is described in Supplemental Data Set 12.

Transient gene expression and cell death assay

Transient expression of ZAR1 and RLCK homologs in *N. benthamiana* was performed by agroinfiltration according to methods described by Bos et al. (2006). Briefly, 4-wk-old *N. benthamiana* plants were infiltrated with *Agrobacterium tumefaciens* strains carrying the binary expression plasmids. *A. tumefaciens* suspensions were prepared in infiltration buffer (10 mM MES, 10 mM $MgCl_2$, and 150 μM acetosyringone, pH 5.6) and were adjusted to appropriate OD_{600} (Supplemental Data Set 12). Macroscopic cell death phenotypes were scored according to Supplemental Fig. 20 and statistical differences among the samples were analyzed with Tukey's HSD test (Supplemental Data Set 13).

RNA-seq data analyses

Public RNA-seq reads, which were previously obtained with Illumina HiSeq 2000 (Chaw et al. 2019), were used to analyze expression profiles of CmZAR1 and ZAR1-CIN genes in the stout camphor tree (Accession Numbers: SRR7416905, SRR7416906, SRR7416908, SRR7416909, SRR7416910, SRR7416911, and SRR7416918). Reads were mapped to the stout camphor genome assembly (GenBank assembly accession GCA_003546025.1) using HISAT2 (Kim et al. 2019) and transformed into a Transcripts Per Million (TPM) value according to Li et al. (2010). TPM values were visualized by the heatmap. The heatmap was colored by 8 ranges (0, 0–5, 5–20, 20–40, 40–60, 60–80, 80–100, and >100) of TPM values.

RNA-seq reads (Accession Numbers: SRR1538828, SRR1538829, SRR1538848, SRR1538903, SRR1538904, SRR1538905, SRR1538928, SRR1538929, SRR1538930, SRR1538931, SRR1538932, SRR1538933, SRR1050891, SRR1050897, SRR1050892, SRR1050898, and SRR3629840) were used to analyze MeZAR1-ID expression in cassava. RNA-seq reads were filtered and trimmed using fastp (Chen et al. 2018). The quality-trimmed reads were mapped to the cassava genome assembly (GenBank assembly accession GCF_001659605.2) using HISAT2 (Kim et al. 2019). Mapped reads were analyzed using Integrative Genomics Viewer (Robinson et al. 2011).

Accession numbers

DNA sequence data used in this study can be found from reference genome or GenBank/EMBL databases with accession numbers listed in Supplemental Data Sets 1, 4, 8, and 9.

Acknowledgments

We are thankful to several colleagues for discussions and ideas. We thank Sebastian Schornack (Sainsbury Laboratory, University of Cambridge, Cambridge, UK) and Kodai Honda (Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University) for valuable comments on this paper. We thank the Prime Minister of the United Kingdom for announcing a stay-at-home order on March 23, 2020.

Author contributions

Conceptualization was done by H.A. and S.K.; data curation was done by H.A., T.S., J.K., H.P., J.L.G.H., and A.M.; formal analysis was done by H.A., T.S., and A.M.; investigation was done

by H.A., T.S., and A.M.; methodology was done by H.A., T.S., J.K., and A.M.; the resources were accumulated by H.A., T.S., J.K., H.P., Y.U., and M.S.; supervision was done by H.A., A.M., and S.K.; funding acquisition was done by S.K.; project administration was done by S.K.; writing the initial draft was done by H.A., J.K., and S.K.; editing was done by H.A., T.S., J.K., J.L.G.H., and S.K.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Arabidopsis ZAR1 is the most conserved CC-NLR across angiosperms, which supports Fig. 1.

Supplemental Figure S2. NbZAR1 is highly conserved across angiosperms, which supports Fig. 1.

Supplemental Figure S3. Sequence alignment of full-length ZAR1 ortholog proteins across angiosperms, which supports Fig. 2.

Supplemental Figure S4. Schematic representation of the intragenomic relationship at ZAR1 loci across angiosperm genomes, which supports Fig. 2.

Supplemental Figure S5. E11 on the glutamate ring inside of the Arabidopsis ZAR1 resistosome is conserved across the orthologs, which supports Fig. 4.

Supplemental Figure S6. Sequence alignment of full-length ZRK proteins across angiosperms, which supports Fig. 5.

Supplemental Figure S7. Heterologous expression of ZAR1 orthologs from flowering plant species in *N. benthamiana*, which supports Fig. 6.

Supplemental Figure S8. *C. esculenta* ZRK1 does not alter autoimmune cell death by *C. esculenta* ZAR1 in *N. benthamiana*, which supports Fig. 6.

Supplemental Figure S9. Four *A. coerulea* ZRKs and 7 *C. micranthum* ZRKs positively regulate AcZAR1 and CmZAR1 autoactive cell death in *N. benthamiana*, which supports Fig. 6.

Supplemental Figure S10. Cell death assay by coexpressing Arabidopsis RKS1 with *A. coerulea* ZAR1 and *C. micranthum* ZAR1 in *N. benthamiana*, which supports Fig. 6.

Supplemental Figure S11. Silencing of *JIM2* does not affect CeZAR1 autoactive cell death in *N. benthamiana*, which supports Fig. 6.

Supplemental Figure S12. Cassava ZAR1-ID is transcribed as a splicing variant from a single locus on the genome, which supports Fig. 7.

Supplemental Figure S13. Trx domain integration occurred in 2 independent rosid ZAR1 subclades, which supports Fig. 7.

Supplemental Figure S14. Integrated Trx domains show high sequence similarity to ZAR1-linked PLP3a gene in Arabidopsis, which supports Fig. 7.

Supplemental Figure S15. ZAR1-SUB gene is distributed across eudicots, which supports Fig. 8.

Supplemental Figure S16. Sequence alignment of full-length ZAR1 and ZAR1-SUB proteins, which supports Fig. 9.

Supplemental Figure S17. ZAR1 and the sister subclade NLRs display different conserved surfaces on the resistosome structure, which supports Fig. 9.

Supplemental Figure S18. ZAR1-CIN gene cluster occurs in the *C. micranthum* genome, which supports Fig. 8.

Supplemental Figure S19. Sequence alignment of full-length ZAR1 and ZAR1-CIN proteins, which supports Fig. 9.

Supplemental Figure S20. Representative images for scoring cell death intensity as an HR index, which supports Fig. 6.

Supplemental Table S1. List of MEME motifs predicted from ZAR1 in angiosperms.

Supplemental Table S2. List of MEME motifs predicted from ZAR1-SUB.

Supplemental Table S3. Comparison of MEME motifs between ZAR1-SUB and ZAR1.

Supplemental Table S4. List of MEME motifs predicted from ZAR1-CIN.

Supplemental Table S5. Comparison of MEME motifs between ZAR1-CIN and ZAR1.

Supplemental Data Set 1. List of ZAR1 in angiosperms. "NF" means "not found."

Supplemental Data Set 2. List of plant species with the number of ZAR1, ZAR1-SUB, and ZAR1-CIN genes.

Supplemental Data Set 3. List of the closest NLR genes to ZAR1 locus.

Supplemental Data Set 4. Genome loci of ZAR1 and ZRK genes. "NA" means "not acquired."

Supplemental Data Set 5. List of genes genetically linked to ZAR1 in eudicots. "NF" means "not found."

Supplemental Data Set 6. List of ZAR1-SUB. "NF" means "not found."

Supplemental Data Set 7. List of ZAR1-CIN. "NF" means "not found."

Supplemental Data Set 8. Reference genome databases used for NLR annotation with NLR-parser.

Supplemental Data Set 9. The ConSurf conservation score among ZAR1 proteins.

Supplemental Data Set 10. The ConSurf conservation score among ZAR1-SUB proteins.

Supplemental Data Set 11. The ConSurf conservation score among ZAR1-CIN proteins.

Supplemental Data Set 12. Plasmid list used in this study.

Supplemental Data Set 13. Summary of Tukey's HSD test results in cell death assay.

Supplemental File 1. Amino acid sequences of full-length ZRKs.

Supplemental File 2. Amino acid alignment file of 35 ZRK in angiosperms.

Supplemental File 3. Amino acid alignment file of 120 ZAR1 in angiosperms.

Supplemental File 4. Amino acid sequences of full-length NLRs used for phylogenetic analysis in Fig. 1B.

Supplemental File 5. Amino acid sequences of full-length NLRs used for phylogenetic analysis in Supplemental Fig. S1.

Supplemental File 6. Amino acid sequences of 120 ZAR1 in angiosperms.

Supplemental File 7. Amino acid sequences of 129 ZAR1-SUB.

Supplemental File 8. Amino acid sequences of 11 ZAR1-CIN.

Supplemental File 9. NLR phylogenetic tree file in Fig. 1B.

Supplemental File 10. NLR phylogenetic tree file in Supplemental Fig. S1.

Supplemental File 11. NLR phylogenetic tree file in Fig. 2.

Supplemental File 12. ZRK phylogenetic tree file in Fig. 5A.

Supplemental File 13. NLR phylogenetic tree file in Fig. 8.

Supplemental File 14. Amino acid alignment file of 129 ZAR1-SUB.

Supplemental File 15. Amino acid alignment file of 11 ZAR1-CIN.

Funding

This work was funded by the Gatsby Charitable Foundation, Biotechnology and Biological Sciences Research Council (BBSRC, UK), and European Research Council (ERC BLASTOFF projects). J.L.G.H. acknowledges support from the South Dakota Agricultural Experiment Station (SD00H605-16). H.A. was funded by the Japan Society for the Promotion of Science (21K20583 and 22K14893) and Japan Science and Technology Agency Precursory Research for Embryonic Science and Technology (JPMJPR21D1).

Conflict of interest statement. S.K. receives funding from industry on NLR biology.

Data availability

Accession numbers for all genes reported and all large-scale data are provided in the manuscript.

References

Adachi H, Contreras MP, Harant A, Wu CH, Derevnina L, Sakai T, Duggan C, Moratto E, Bozkurt TO, Maqbool A, et al. An N-terminal motif in NLR immune receptors is functionally conserved across distantly related plant species. eLife. 2019b:**8**:e49956. https://doi.org/10.7554/eLife.49956

Adachi H, Derevnina L, Kamoun S. NLR singletons, pairs, and networks: evolution, assembly, and regulation of the intracellular immunoreceptor circuitry of plants. Curr Opin Plant Biol. 2019a:**50**:121–131. https://doi.org/10.1016/j.pbi.2019.04.007

Adachi H, Sakai T, Kourelis J, Maqbool A, Kamoun S. Jurassic NLR: conserved and dynamic evolutionary features of the atypically ancient immune receptor ZAR1. bioRxiv. 2020. https://doi.org/10.1101/2020.10.12.333484

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990:**215**(3):403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ashkenazy H, Abadi S, Martz E, et al. Consurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 2016:**44**(W1):W344–W350. https://doi.org/10.1093/nar/gkw408

Baggs E, Dagdas G, Krasileva KV. NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. Curr Opin Plant Biol. 2017:**38**:59–67. https://doi.org/10.1016/j.pbi.2017.04.012

Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. 1994:**2**:28–36.

Baudin M, Hassan JA, Schreiber KJ, Lewis JD. Analysis of the ZAR1 immune complex reveals determinants for immunity and molecular interactions. Plant Physiol. 2017:**174**(4):2038–2053. https://doi.org/10.1104/pp.17.00441

Baudin M, Schreiber KJ, Martin EC, Petrescu AJ, Lewis JD. Structure-function analysis of ZAR1 immune receptor reveals key molecular interactions for activity. Plant J. 2020:**101**(2):352–370. https://doi.org/10.1111/tpj.14547

Bayless AM, Nishimura MT. Enzymatic functions for Toll/interleukin-1 receptor domain proteins in the plant immune system. Front Genet. 2020:**11**:539. https://doi.org/10.3389/fgene.2020.00539

Bi G, Su M, Li N, et al. The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. Cell. 2021:**184**(13):3528–3541.e12. https://doi.org/10.1016/j.cell.2021.05.003

Bos JI, Kanneganti TD, Young C, Cakir C, Huitema E, Win J, Armstrong MR, Birch PR, Kamoun S. The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1- induced cell death in *Nicotiana benthamiana*. Plant J. 2006:**48**(2):165–176. https://doi.org/10.1111/j.1365-313X.2006.02866.x

Burdett H, Bentham AR, Williams SJ, et al. The plant "resistosome": structural insights into immune signaling. Cell Host Microbe. 2019:**26**(2):193–201. https://doi.org/10.1016/j.chom.2019.07.020

Cesari S, Bernoux M, Moncuquet P, Kroj T, Dodds PN. A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. Front Plant Sci. 2014:**5**:606. https://doi.org/10.3389/fpls.2014.00606

Chaw SM, Liu YC, Wu YW, et al. Stout camphor tree genome fills gaps in understanding of flowering plant genome evolution. Nat Plants. 2019:**5**(1):63–73. https://doi.org/10.1038/s41477-018-0337-0

Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018:**34**(17):i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Coiro M, Doyle JA, Hilton J. How deep is the conflict between molecular and fossil evidence on the age of angiosperms? New Phytol. 2019:**223**(1):83–99. https://doi.org/10.1111/nph.15708

Cui D-F, Hou Y, Yin P, Wang X. A Jurassic flower bud from China. Geological Society, London. 2022:**521**:81–93. https://doi.org/10.1144/SP521-2021-122

Dangl JL, Horvath DM, Staskawicz BJ. Pivoting the plant immune system from dissection to deployment. Science. 2013:**341**(6147):746–751. https://doi.org/10.1126/science.1236011

Delaux PM, Hetherington AJ, Coudert Y, et al. Reconstructing trait evolution in plant evo-devo studies. Curr Biol. 2019:**29**(21):R1110–R1118. https://doi.org/10.1016/j.cub.2019.09.044

Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet. 2010:**11**(8):539–548. https://doi.org/10.1038/nrg2812

Eddy SR. Profile hidden Markov models. Bioinformatics. 1998:**14**(9):755–763. https://doi.org/10.1093/bioinformatics/14.9.755

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JD, Patron NJ, Marillonnet S. A golden gate modular cloning toolbox for plants. ACS Synth Biol. 2014:**3**(11):839–843. https://doi.org/10.1021/sb4001504

Feehan JM, Castel B, Bentham AR, Jones JD. Plant NLRs get by with a little help from their friends. Curr Opin Plant Biol. 2020:**56**:99–108. https://doi.org/10.1016/j.pbi.2020.04.006

Filiault DL, Ballerini ES, Mandáková T, et al. The *Aquilegia* genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. ELife. 2018:**7**:e36426. https://doi.org/10.7554/eLife.36426

Fu Q, Diez JB, Pole M, et al. An unexpected noncarpellate epigynous flower from the Jurassic of China. Elife. 2018:**7**:e38827. https://doi.org/10.7554/eLife.38827

Gong Z, Qi J, Hu M, Bi G, Zhou JM, Han GZ. The origin and evolution of a plant resistosome. Plant Cell. 2022:**34**(5):1600–1620. https://doi.org/10.1093/plcell/koac053

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. 2022:**188**(1):70–80. https://doi.org/10.1093/plphys/kiab471

Hu M, Qi J, Bi G, Zhou JM. Bacterial effectors induce oligomerization of immune receptor ZAR1 *in vivo*. Mol Plant. 2020:**13**(5):793–801. https://doi.org/10.1016/j.molp.2020.03.004

Jones P, Binns D, Chang H, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. Interproscan 5: genome-scale protein function classification. Bioinformatics. 2014:**30**(9):1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jones JD, Dangl JL. The plant immune system. Nature. 2006:**444**(7117):323–329. https://doi.org/10.1038/nature05286

Jones JD, Vance RE, Dangl JL. Intracellular innate immune surveillance devices in plants and animals. Science. 2016:**354**(6316):aaf6395. https://doi.org/10.1126/science.aaf6395

Jubic LM, Saile S, Furzer OJ, El Kasmi F, Dangl JL. Help wanted: helper NLRs and plant immune responses. Curr Opin Plant Biol. 2019:**50**:82–94. https://doi.org/10.1016/j.pbi.2019.03.013

Karasov TL, Chae E, Herman JJ, Bergelson J. Mechanisms to mitigate the trade-off between growth and defense. Plant Cell. 2017:**29**(4):666–680. https://doi.org/10.1105/tpc.16.00931

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:**30**(4):772–780. https://doi.org/10.1093/molbev/mst010

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015:**10**(6):845–858. https://doi.org/10.1038/nprot.2015.053

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnol. 2019:**37**(8):907–915. https://doi.org/10.1038/s41587-019-0201-4

Kourelis J, Sakai T, Adachi H, Kamoun S. RefPlantNLR is a comprehensive collection of experimentally validated plant disease resistance proteins from the NLR family. PLoS Biol. 2021:**19**(10):e3001124. https://doi.org/10.1371/journal.pbio.3001124

Kourelis J, van der Hoorn RAL. Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. Plant Cell. 2018:**30**(2):285–299. https://doi.org/10.1105/tpc.17.00579

Kroj T, Chanclud E, Michel-Romiti C, Grand X, Morel JB. Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. New Phytol. 2016:**210**(2):618–626. https://doi.org/10.1111/nph.13869

Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016:**33**(7):1870–1874. https://doi.org/10.1093/molbev/msw054

Laflamme B, Dillon MM, Martel A, Almeida RND, Desveaux D, Guttman DS. The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. Science. 2020:**367**(6479):763–768. https://doi.org/10.1126/science.aax4079

Lee RRQ, Chae E. Variation patterns of NLR clusters in *Arabidopsis thaliana* genomes. Plant Communications. 2020:**1**(4):100089. https://doi.org/10.1016/j.xplc.2020.100089

Lee H-Y, Mang H, Choi E-H, Seo Y-E, Kim M-S, Oh S, Kim S-B, Choi D. Genome-wide functional analysis of hot pepper immune receptors reveals an autonomous NLR cluster in seed plants. New Phytol. 2021:**229**(1):532–547. https://doi.org/10.1111/nph.16878

Lewis JD, Lee AH, Hassan JA, 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. 2013:**110**(46):18722–18727. https://doi.org/10.1073/pnas.1315520110

Lewis JD, Wu R, Guttman DS, Desveaux D. Allele-specific virulence attenuation of the *Pseudomonas syringae* HopZ1a type III effector via the *Arabidopsis* ZAR1 resistance protein. PLoS Genet. 2010:**6**(4):e1000894. https://doi.org/10.1371/journal.pgen.1000894

Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics. 2010:**26**(4):493–500. https://doi.org/10.1093/bioinformatics/btp692

Lin X, Zhang Y, Kuang H, Chen J. Frequent loss of lineages and deficient duplications accounted for low copy number of disease resistance genes in Cucurbitaceae. BMC Genomics. 2013:**14**:335. https://doi.org/10.1186/1471-2164-14-335

Ma S, Lapin D, Liu L, et al. Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. Science. 2020:**370**(6521):eabe3069. https://doi.org/10.1126/science.abe3069

Martin R, Qi T, Zhang H, Liu F, King M, Toth C, Nogales E, Staskawicz BJ. Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. Science. 2020:**370**(6521):eabd9993. https://doi.org/10.1126/science.abd9993

Mermigka G, Amprazi M, Mentzelopoulou A, Amartolou A, Sarris PF. Plant and animal innate immunity complexes: fighting different enemies with similar weapons. Trends Plant Sci. 2020:**25**(1):80–91. https://doi.org/10.1016/j.tplants.2019.09.008

Michelmore RW, Meyers BC. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. 1998:**8**(11):1113–1130. https://doi.org/10.1101/gr.8.11.1113

Pai H, Kourelis J, Adachi H, Kamoun S. ZED1-related kinases (ZRKs) from *Aquilegia coerulea* and *Cinnamomum micranthum* do not trigger autoactive cell death in *Nicotiana benthamiana*. Zenodo. 2023. https://doi.org/10.5281/zenodo.7844723

Prigozhin DM, Krasileva KV. Analysis of intraspecies diversity reveals a subset of highly variable plant immune receptors and predicts their binding sites. Plant Cell. 2021:**33**(4):998–1015. https://doi.org/10.1093/plcell/koab013

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011:**29**(1):24–26. https://doi.org/10.1038/nbt.1754

Sarris PF, Cevik V, Dagdas G, Jones JD, Krasileva KV. Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. BMC Biol. 2016:**14**:8. https://doi.org/10.1186/s12915-016-0228-7

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. 2019:**221**(2):1001–1009. https://doi.org/10.1111/nph.15411

Seong K, Seo E, Witek K, Li M, Staskawicz B. Evolution of NLR resistance genes with noncanonical N-terminal domains in wild tomato species. New Phytol. 2020:**227**(5):1530–1543. https://doi.org/10.1111/nph.16628

Seto D, Koulena N, Lo T, Menna A, Guttman DS, Desveaux D. Expanded type III effector recognition by the ZAR1 NLR protein using ZED1-related kinases. Nat Plants. 2017:**3**:17027. https://doi.org/10.1038/nplants.2017.27

Shao ZQ, Xue JY, Wu P, 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. 2016:**170**(4):2095–2109. https://doi.org/10.1104/pp.15.01487

Smith SA, Brown JW. Constructing a broadly inclusive seed plant phylogeny. Am J Bot. 2018:**105**(3):302–314. https://doi.org/10.1002/ajb2.1019

Stam R, Silva-Arias GA, Tellier A. Subsets of NLR genes show differential signatures of adaptation during colonization of new habitats. New Phytol. 2019:**224**(1):367–379. https://doi.org/10.1111/nph.16017

Steuernagel B, Jupe F, Witek K, Jones JD, Wulff BB. NLR-parser: rapid annotation of plant NLR complements. Bioinformatics. 2015:**31**(10): 1665–1667. https://doi.org/10.1093/bioinformatics/btv005

Sukumaran J, Holder MT. Dendropy: a Python library for phylogenetic computing. Bioinformatics. 2010:**26**(12):1569–1571. https://doi.org/10.1093/bioinformatics/btq228

Tamborski J, Krasileva KV. Evolution of plant NLRs: from natural history to precise modifications. Annu Rev Plant Biol. 2020:**71**:355–378. https://doi.org/10.1146/annurev-arplant-081519-035901

Uehling J, Deveau A, Paoletti M. Do fungi have an innate immune response? An NLR-based comparison to plant and animal immune systems. PLoS Pathog. 2017:**13**(10):e1006578. https://doi.org/10.1371/journal.ppat.1006578

Van de Weyer AL, Monteiro F, Furzer OJ, et al. A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. Cell. 2019:**178**(5):1260–1272.e14. https://doi.org/10.1016/j.cell.2019.07.038

Wang J, Chai J. Structural insights into the plant immune receptors PRRs and NLRs. Plant Physiol. 2020:**182**(4):1566–1581. https://doi.org/10.1104/pp.19.01252

Wang J, Hu M, Wang J, Qi J, Han Z, Wang G, Qi Y, Wang H-W, Zhou J-M, Chai J. Reconstitution and structure of a plant NLR resistosome conferring immunity. Science. 2019b:**364**(6435):eaav5870. https://doi.org/10.1126/science.aav5870

Wang G, Roux B, Feng F, 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. 2015:**18**(3): 285–295. https://doi.org/10.1016/j.chom.2015.08.004

Wang J, Wang J, Hu M, Shan W, Qi J, Wang G, Han Z, Qi Y, Gao N, Wang H-W, et al. Ligand-triggered allosteric ADP release primes a plant NLR complex. Science. 2019a:**364**(6435):eaav5868. https://doi.org/10.1126/science.aav5868

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS One. 2011:**6**(2):e16765. https://doi.org/10.1371/journal.pone.0016765

Wu CH, Abd-El-Haliem A, Bozkurt TO, et al. NLR network mediates immunity to diverse plant pathogens. Proc Natl Acad Sci U S A. 2017:**114**(30):8113–8118. https://doi.org/10.1073/pnas.1702041114

Wu CH, Derevnina L, Kamoun S. Receptor networks underpin plant immunity. Science. 2018:**360**(6395):1300–1301. https://doi.org/10.1126/science.aat2623

Wu CH, Krasileva KV, Banfield MJ, Terauchi R, Kamoun S. The "sensor domains" of plant NLR proteins: more than decoys? Front Plant Sci. 2015:**6**:134. https://doi.org/10.3389/fpls.2015.00134

Xiong Y, Han Z, Chai J. Resistosome and inflammasome: platforms mediating innate immunity. Curr Opin Plant Biol. 2020:**56**:47–55. https://doi.org/10.1016/j.pbi.2020.03.010

Zhou JM, Zhang Y. Plant immunity: danger perception and signaling. Cell. 2020:**181**(5):978–989. https://doi.org/10.1016/j.cell.2020.04.028