Plant Receptor-like proteins (RLPs): Structural features enabling versatile immune recognition

Simon Snoeck[1], Anthony GK Garcia[1], Adam D Steinbrenner*[1]

1, University of Washington Department of Biology, Box 351800, Seattle WA 98195

*corresponding author, astein10@uw.edu

Abstract

Plant immune recognition of pests and pathogens relies on a germline-encoded repertoire of innate immune receptors. In recent years, many examples of cell surface receptors in the large gene family of receptor-like proteins (RLPs) have been shown to govern recognition specificity. RLPs lack the canonical intracellular kinase domain associated with receptor kinases, but instead transduce immune signaling through interactions with co-receptor and adaptor kinases including SOBIR1. We review recognition and signaling functions mediated by subdomains of leucine-rich repeat (LRR) type RLPs. LRR-RLP extracellular domains, transmembrane motifs, and intracellular tails can mediate ligand binding, co-receptor recruitment, and immune signaling specificity. A recently reported cryo-EM structure of an LRR-RLP, RXEG1, now links domain architecture and sequence features of LRR-RLPs to key molecular functions. Finally, we propose a new motif-based classification of LRR-RLPs based on shared sequence features in the island domain (ID). In both dicot and monocot species, receptors encoding the Y-x_8-KG and the newly defined K-x_5-Y motif in their IDs correspond with conserved, sister clades of LRR-RLP genes. We propose that the conservation of distinct Y-x_8-KG and K-x_5-Y groups of LRR-RLPs implies that the motif is a useful classifier of immunity-related LRR-RLPs, distinguishing them from receptors with similar domain architecture but roles in growth and development. Deeply conserved structural features of LRR-RLPs likely provide a versatile platform for diversification of receptor repertoires and recognition specificities.

1. Introduction

Plant health is mediated by successful immune recognition and response against pests and pathogens. Initial recognition is achieved by innate immune receptors which sense non-self or modified-self patterns. Receptors are divided into two categories based on their localization and types of detected patterns [1,2]. A first set of intracellular receptors in the NOD-like receptor (NLR) family monitors pathogen effector proteins [3,4]. A second set of receptors, termed pattern recognition receptors (PRRs), resides at the plasma membrane and monitors extracellular signals, termed pathogen/pest-associated molecular patterns (PAMPs) [5,6].

Canonical PRRs in the receptor kinase (RK) family, such as Flagellin Sensing 2 (FLS2), contain two distinct domains to carry out dual functions: recognition and signal transduction. For recognition, many different extracellular domains can bind PAMPs, but highly variable, proteinaceous PAMPs are typically detected by receptors containing a leucine-rich repeat (LRR) ectodomain [7]. For signal transduction, most PRRs encode an intracellular kinase domain and associate with key small-ectodomain co-receptors in the SERK family (e.g. BAK1), part of a broad network of homotypic (LRR-LRR) protein-protein interaction network [8,9]. In addition to immune sensing functions, RK have diverse roles in abiotic sensing and plant development [10,11].

In contrast to RKs, a second set of PRRs termed Receptor-like Proteins (RLPs) do not encode this simple dual functionality in their domain architecture. RLPs contain an ectodomain, transmembrane, and short cytoplasmic tail, but not an intracellular kinase domain for signal transduction [12,13]. RLPs were among the first disease resistance genes to be cloned [14,15], but mechanisms for recognition and signaling are still far less well understood than for NLRs and RKs. This review focuses on the largest subfamily of RLPs in plants, those containing extracellular leucine-rich repeat (LRRs). Beyond a few examples of LRR-RLPs with roles in development, most LRR-RLPs mediate the detection of various threats such as fungi, parasitic weeds and herbivores [16–20]. LRR-RLPs and their respective ligands have been characterized in diverse species such as Arabidopsis, tomato, wild tobacco, canola and cowpea and detect diverse molecular patterns from short AA peptides [16,17] to specific features of protein structures, for example the active site groove of fungal XEG1 and the tertiary fold of bacterial translation initiation factor-1 (IF1) [21,22]. A recent breakthrough description of the cryo-EM structure of the LRR-RLP RXEG1, as well as recent functional studies of RLP42 and INR function [21,23,24], motivates a new overview of LRR-RLP structure and function.

A major goal in the field of plant immunobiology is to identify agronomically useful receptors [25,26]. Growing information on immune-related LRR-RLPs now allows us to define sequence motifs of immune-related LRR-RLPs. Unlike RKs and NLRs, which have relatively conserved signaling domains for phylogenetic classification, LRR-RLPs lack a discrete signaling domain, and there is not yet a unified nomenclature for describing LRR-RLP subfamilies [12,13,27]. However, new structural data and comparative genomics reveal that LRR-RLPs involved in PAMP recognition differ from receptors involved in plant development, distinguished by key differences in the functionally critical island domain (ID). Here we define two clades of immunity-related LRR-RLPs, those with Y-x_8-KG and K-x_5-Y IDs. We hope that this nomenclature will facilitate discussion of immunity-related LRR-RLPs in future genomic and functional analyses.

2. LRR-RLP Domain Architecture

The extracellular LRR domain of plant transmembrane receptors forms a twisted or superhelical assembly with a β-sheet forming the inner surface of the LRRs [10]. Seven regions, domains A through G, were originally defined in the sequence of Cf-9, the first characterized LRR-RLP [14]. The B- and D- domain are LRR capping domains stabilized by disulfide bonds, which likely shield the hydrophobic core of the LRR domain at respectively the N- and C-terminus [28]. The LRR or C-domain was later divided into three subdomains (C1-C3), with the non-LRR-containing C2 subdomain dividing the fourth and the fifth to last C-terminal LRR [29]. The C2 subdomain was later alternatively named the intervening motif, island region/domain, or carboxy terminal (CT) loopout domain [21,29]. Here we use the term island domain (ID) as in Sun et al. (2022). Following the C1-C3 ID+LRRs, the E-domain is also called the acidic domain due to the strong presence of negatively charged AAs. A hydrophobic transmembrane helix (F-domain) and an unstructured intracellular and mostly basic charged cytoplasmic tail (G-domain) follow at the C-terminus of LRR-RLPs [14].

3. Keeping host plants “in the loop” – A ligand-binding role for the long-neglected N-terminal loopouts in domain B or C1.

The cryoEM structure of RXEG1 recently revealed that “loopouts” in the classical LRR solenoid structure are not restricted to the ID. An 11-AA loopout structure embedded at the N-terminal (NT) of the C1 LRR domain was shown to be crucial for RXEG1 function as it is involved in the interaction with its ligand XEG1 as well as competitive inhibition of XEG1 endoglucanase [21]. An RXEG1 mutant with the middle of the NT loopout removed (RXEG1$^{\Delta 90-95}$) impaired cell death response after exposure to XEG1 and nearly abolished competitive inhibition of XEG1 [21]. This breakthrough discovery adds to the theme of ligand-binding roles for loopouts in LRR domains of both plant NLRs and human Toll-like receptors [30,31].

Here we investigated all identified LRR-RLPs with a known ligand and defined their NT loopout domains by using a standardized cut-off based on an alignment and visual validation based on AlphaFold2 (AF2) structure predictions and the cryoEM structure of RXEG1 (Fig. 1, Suppl. Table 1, Supplemental files 1-3). 17 of 21 investigated LRR-RLPs contain one or more NT loopouts, and these vary in size from 2-19 AA (Fig. 1B). NT loopouts are embedded within either (1) the NT capping domain (B-domain) which is characterized by 2-6 Cys residues, or (2) the N-terminal LRRs, as with RXEG1. Intriguingly, although some LRR-RLPs share conserved portions of their NT loopout sequence, none contain exactly the same NT loopout, even among closely related LRR-RLPs such as Cf-4, Cf-9 and Hcr9-4E. Cys residues within the B-domain were earlier predicted to stabilize the NT cap by forming disulfide bonds by Matsushima et al. [32] ,and single AA substitutions to Ala of two strongly conserved Cys residues resulted in reduction of Cf-9 function [28]. Additionally, AF2 structure predictions suggest that 9 NT loopouts form secondary structure, but that secondary structure is missing in other NT loopouts. Additional LRR-RLP structures are needed to validate AF2 predictions and to draw firm conclusions regarding shared features of loopout structure.

Although the NT loopout was not defined prior to the elucidation of the RXEG1 structure, three additional studies had previously provided functional insights into the role of the NT loopout. First, Cf-9 has an extended version of the Cf-4 NT loopout. Chimeric Cf-4 / Cf-9 receptors revealed that an HR response to Avr4 is weakened by swapping the Cf-9 extended version of the NT loopout in a Cf-4 backbone. In contrast, Cf-9 response to Avr9 is not weakened by swapping in the relatively short Cf-4 NT loopout, implying that Cf-9 specificity for Avr9 resides only in the LRRs [33,34]. Second, a Cys to Ala substitution within the NT loopout of Cf-9 resulted in a slightly reduced activity of Cf-9. Third, an ancestral sequence reconstruction approach applied to functional and non-functional homologues of the LRR-RLP INR pinpointed a single Tyr residue in the NT loopout that is strongly conserved within all identified INR homologues from 11 legume species. An ancestral INR mutant with an AA substitution at this site resulted in a relatively lower ROS burst [24]. In summary, there is evidence for a crucial role of the NT loopout for several LRR-RLPs, and the observed variation across NT loopouts suggests a function in ligand specificity. However, the NT loopout may not be required for ligand binding in all cases, as the predicted protein structures of Cf-2, Cf-5, RLP23 and RLP42 do not contain a clear NT loopout.

4. Recognition roles of C1 LRRs

A series of LRRs make up the largest LRR-RLP domain, C1, and abundant evidence points to a role for C1 in recognition specificity. Although Cf-4 and Cf-9 have a high AA similarity of 91.68%, Cf-9 contains two additional LRR (LRR 11-12) in its C1 domain. The addition or deletion of this region to respectively Cf-4 and Cf-9 abolished the function of both receptors [33,34], suggesting the importance of the relative distances of specific AA residues for at least Cf-4 receptor function. Chimeric receptors and PCR-mediated gene shuffling based on Cf-4 and Cf-9 revealed the importance of divergent AA residues within the LRRs 11-14 and LRR 10-16 respectively for Cf-4 and Cf-9 [33–35]. Considering Cf-9, AA L457 in LRR 16 was shown to be essential for Cf-9 function as well as the presence of LRR 11-12, which are absent in Cf-4 [33,34]. Within Cf-4 LRR 11-14, Van der Hoorn *et al.* were able to pinpoint three specific residues W389, G411, and F457 which are essential for full Avr4 HR response [34]. Moreover, introduction of these in a Cf-9 backbone in combination with the deletion of LRR 11-12 was sufficient to gain Avr4 HR response. Using a similar approach for Cf-9 and Cf-9B, 4 specific residues in the LRR 13-15 of Cf-9 were identified as specificity determining residues. When incorporated in a Cf-9B backbone, these amino acids were sufficient to gain Avr9 HR response [35].

Beyond the *Solanum* LRR-RLPs, truncations of the *Arabidopsis* RLP23 C1 domain revealed that the region spanning LRR1 and LRR3 is crucial for ligand binding and RLP23 function [36], and chimeric receptors of RLP42 and RLP40 implicated the importance of the 12 N-terminal LRRs of RLP42 [23]. Many individual RLP42→RLP40 substitutions located within these LRRs were also sufficient to abolish function. Cowpea INR and INR-like 1 chimeric receptors similarly revealed the necessity of the C1 domain for INR function, and two AA residues in LRR 14 (H404 and R406) were found to be crucial for INR function through an ancestral sequence reconstruction approach [24]. Finally, the RXEG1 cryoEM structure indicates potential interaction between two C1 domain residues, C247 and N674, with the XEG1 ligand, although this was not further analyzed [21].

5. Island domain (ID) roles in ligand binding and BAK1 co-receptor association

Following C1 in the LRR-RLP domain architecture, a non-LRR loopout termed the island domain (ID) is present in most LRR-RLPs – and is indeed characteristic of PAMP-recognizing LRR-RLPs as we discuss below. The ID is always positioned N-terminal to the last set of exactly four LRRs, termed C3. Genetic studies of LRR-RLPs RLP42, INR and RXEG1 provide evidence for the importance of the ID for LRR-RLP function [21,23,24,37]. Chimeric receptors with a non-functional paralog demonstrated a requirement of the ID for maintaining the function of INR [24]. Similarly for Ve1 and RLP42, chimeric receptors showed the requirement of the ID for specificity (Fradin 2014, Zhang 2021). Moreover, a mutant with an AA substitution in the ID (E696K) had a reduced ligand binding and impaired RLP42 function [23]. Finally, the crystal structure of RXEG1 revealed multiple potential interactions of the ID with its ligand XEG1, and an ID mutant of RXEG1 (Q786A/K791A) had impaired cell death activation (Sun et al., 2022). Consistent with a role for the ID in co-receptor association but not recognition specificity, both pairs of Cf genes, Cf-2 and Cf-5, and Cf-4 and Cf-9, share the same ID sequence but respond to different effectors [34,38–40].

The RXEG1 structure now indicates that the role of the ID extends to co-receptor association. LRR-RLPs lack an intracellular kinase domain and several of them have been shown to associate with an LRR-RLK, SUPPRESSOR OF BIR1-1 (SOBIR1), for signaling activation [41,42], an aspect of signaling that we discuss in Box 1. Activation of the RLP-SOBIR1 complex was shown to require SERK co-receptors such as BAK1 [43–45]. The RXEG1 structure now indicates that flexibility or LRR-RLP IDs upon ligand binding mediates inducible association with BAK1 [21]. The combination of two structures, the RXEG1 ligand-free structure and RXEG1 in complex with BAK1 and XEG1, reveals that BAK1 contacts the last four LRRs of RXEG1 and the ID but not the ligand XEG1. RXEG1 mutations at contact points either in the ID domain (W806A and K807A) or at residue in the C3 domain, E896A, abolished the interaction with BAK1 *in vivo*. Intriguingly, K807 and E869 are strongly conserved within all LRR-RLPs, as is another BAK1-interacting AA in the C3 domain, RXEG1 Q913 (see alignment positions 1356, 1419, and 1462 in Supplementary file 2). RXEG K807 is also part of the highly conserved K-x_5-Y motif (see section 7). Thus structural data and conservation across LRR-RLPs suggests that these residues are broadly critical for BAK1 complex formation. Since ligand-inducible association of RLPs with SERK co-receptors has been shown for many LRR-RLPs [16,17,20,46,47], a conserved functional role for the ID is likely.

6. ID sequence motifs define two classes of LRR-RLPs: K-x_5-Y and Y-x_8-KG

Similarly to our approach to analyze NT-loopouts, we compiled ID sequences of the same set of immunity-related LRR-RLPs with defined ligands (Supplemental File 1). To define IDs we used conserved cut-offs in an alignment, which were subsequently validated with AF2 structure predictions (Supplemental Table 1). IDs were earlier defined as long loop structures inserted

between two consecutive LRR repeats which may fold into domain-like structures that harbor secondary structural elements. Intriguingly, many of these IDs are considerably larger as initially reported as they contain AAs which are traditionally identified as parts of the LRRs immediately flanking the ID. The cryoEM structure of RXEG1 and AF2 predictions of other LRR-RLPs visualize the impact of these extra AAs, which give additional structure to the C-terminal loopouts by forming beta-strands (Fig. 1A). All LRR-RLP IDs predicted by AF2 have 2-4 beta-strands. Secondary structure elements in IDs were also observed in the crystal structures of the LRR-RLKs BRI1, PSKR, and RPK2, although structure and function of RK IDs may differ from that of LRR-RLPs [10]. IDs analyzed here have an average ID size of 44 AA, with large variation between receptors. For example, ReMAX (RLP1), I gene, and CuRe1 have a relatively long loopout domain with respectively 63, 64 and 65 AAs. In contrast, the ID of Cf-2 and Cf-5 only consists of 33 AAs.

Using the collection of ID sequences, we searched for conserved motifs (Fig. 1C). Two groups of LRR-RLPs were immediately apparent from two alternative sets of Tyr and/or Gly residues with consistent spacing around the same conserved Lys. We term these motifs Y-x_8-KG or K-x_5-Y. Fritz-Laylin (2005) previously noted the Y-x_8-KG motif in multiple IDs in both Arabidopsis and rice LRR-RLP genes. We also observed this motif in 13/21 investigated LRR-RLPs, whereas RLP30 contains the C-terminal residues KG but lacks the N-terminal Y. An alternative motif, K-x_5-Y, which can be found in 7/21 of the investigated LRR-RLPs (Fig. 1C). In RXEG1, a K-x_5-Y type LRR-RLP, the conserved Lys is found at position 807 and mediates BAK1 interaction as described above. No structure of a Y-x_8-KG LRR-RLP is yet available. In summary, characteristic and highly conserved residues appear to be consistently spaced around a conserved, LRR-RLP specific Lys used for co-receptor recruitment.

Categorization of LRR-RLPs by the presence of the alternative K-x_5-Y or Y-x_8-KG motifs in the ID provides a new method for analyzing diversity and evolution of the gene family. LRR-RLPs containing either motif cluster in phylogenetic analyses into large clades of varying gene number in 4 analyzed species (Fig. 2). For Arabidopsis, the two clades contain the majority of previously annotated LRR-RLPs, but notably not RLP17 (TMM), RLP10 (CLV2), or RLP44, consistent with previous phylogenetic analysis [13]. Intriguingly, representative species from both eudicots and monocots contain members of both K-x_5-Y and Y-x_8-KG clades. LRR-RLPs which contain mixed or neither motif are rare and often lack transmembrane helices. The number of genes K-x_5-Y and Y-x_8-KG clades seems to vary independently, possibly reflecting cycles of expansion and deletion. The conservation of two distinct classes of LRR-RLP from monocot to dicot species suggests distinct functional roles for K-x_5-Y and Y-x_8-KG receptors. Previous analyses categorized LRR-RLPs by gene expression patterns (pathogen responsive), by coarse phylogenetic patterns (derived, basal, or species-specific LRR-RLPs), or by relatedness to characterized LRR-RLPs (e.g. Cf-like) [27,29,48,49]. Based on the analysis described above, we propose to instead classify LRR-RLPs as K-x_5-Y type LRR-RLPs, Y-x_8-KG type LRR-RLPs, or outgroup LRR-RLPs lacking these two motifs**.**

K-x_5-Y and Y-x_8-KG motifs appear to be specific to PRRs and are absent in LRR-RLPs mediating developmental phenotypes. For example, TMM, RLP44, and CLV2 mediate stomatal patterning, brassinosteroid signaling, and meristem maintenance. RLP44 lacks an ID, and the TMM and

CLV2 IDs lack either K-x_5-Y and Y-x_8-KG motif. Consistent with this difference in ID motif, molecular mechanisms of developmental LRR-RLP function differ from immunity-related LRR-RLPs. In contrast to the critical role of the ID of various immunity-related LRR-RLPs, a CLV2 ID deletion was still functional as it was able to rescue a *clv2-3* mutant [50]. Structural data also support a highly distinct functional mechanism for the developmental LRR-RLP TMM [51]. The TMM ectodomain associates with the RK Erecta (ER) and alters specificity against EPF peptide ligands, but does not directly bind EPF.

Evolutionary patterns of developmental RLPs are also distinct from immunity related LRR-RLPs. TMM and RLP44 are broadly conserved as single orthologs in distantly related species [52] and not broadly pathogen inducible [27]. In contrast, immunity-related LRR-RLPs fall in large, species specific groups [48,49], consistent with lineage-specific expansions of LRR-RLP repertoires. Finally, K-x_5-Y and Y-x_8-KG LRR-RLPs show high intraspecies variation. Among Arabidopsis LRR-RLPs for example, all receptors with pangenomic variation classified by Pruitt et al. [53] as either “complex” and “presence/absence” fall within the K-x_5-Y and Y-x_8-KG clades, consistent with complex selective pressures leading to locus-level natural variation. Thus, evolutionary differences compared to non-immunity related receptors further support a distinct classification for the K-x_5-Y and Y-x_8-KG type LRR-RLPs

Functional differences between K-x_5-Y and Y-x_8-KG clades are not yet explored. Examples of receptors from both categories can mediate cell death, ROS, and/or ethylene responses against diverse PAMPs. Divergent functions may become clear as LRR-RLP signaling mechanisms are explored further. Interestingly, specific mechanistic links between LRR-RLPs and intracellular NLR signaling machinery have recently been defined, providing a potential mechanism for different classes of LRR-RLP to activate distinct defense responses [53–55]. For example, cell death for Solanaceae-specific RLPs Cf-4 and Cf-9 is mediated by NRC3 [55], but it is not clear whether NLR dependence extends to LRR-RLPs from species outside the Solanaceae, which lack the NRC clade of NLRs. Therefore a critical research priority is to understand specificity in LRR-RLP signaling pathways leading to NLR-dependent and NLR-independent responses.

7. C-terminal capping domain D

Two Cys residues delimit the boundary of the C and D domains of LRR-RLPs [56], with the second Cys strongly conserved within the C-terminal capping domain D. Van der Hoorn 2005 reported earlier the following motif for domain D: GNxGLCGxPLx_3C. However, our LRR-RLP alignment suggests that only Nx_2LxGx_6C is strongly conserved. Intriguingly, using the same strategy as discussed above for investigating the importance of Cys residues in the B-domain, a single AA substitution to Ala of the strongly conserved Cys residue in domain D did not impact Cf-9 function [28].

8. The acidic or juxtamembrane domain E and transmembrane domain F (Gx_3G) are major players in SOBIR1 association.

A final set of domains C-terminal to LRRs and capping domains show LRR-RLP specific features. A set of acidic residues, domain E, is highly conserved across LRR-RLPs and may mediate interactions with SOBIR1 (see Box 1). The single-pass transmembrane helix contains an LRR-RLP specific motif, Gx_3G [57], and is termed domain F.

Evidence from chimeric and truncated LRR-RLPs have indicated a key role for the C-terminal domains. First, Ve1 still associates with SOBIR1 when it is N-terminally truncated to the second to last LRR of its C2-domain [37]. Second, a chimeric receptor with domain A-E of LRR-RLK EFR and domain F and G (transmembrane cytoplasmic tail) of the LRR-RLP Cf-9 was shown to associate with SOBIR1, in contrast to EFR itself [58]. Moreover, similar to EFR, this chimeric receptor was able to respond to the EFR ligand elf18C, and silencing of NtSOBIR1/SOBIR1-like severely compromised the cell death response. Third, truncated RLP23 receptors provide evidence that domain E is required for full-strength association with SOBIR1 [36].

Specific mutations in the Gx_3G also support a key role in SOBIR1 association. Gx_3G motifs were earlier characterized as facilitators of protein–protein interactions in animal receptor tyrosine kinases or membrane channels [59], and the majority of LRR-RLPs contain one or several tandemly arranged Gx_3G motifs [56,57]. Of the 19 immunity-related LRR-RLPs analyzed here, 14 have at least one Gx_3G motif. The exceptions contain either degenerate motifs: RXEG1 and EIX2 contain $\underline{G}x_6\underline{G}x_2\underline{G}$, and CuRE1 contains G_5Gx_4G, or lack any of the conserved Gly residues in their predicted transmembrane domains as for I gene and ReMAX (Suppl. Fig 2) [18,42,60–62]. Hence, the strict Gx_3G motif is dispensable and alternative motives must contribute to SOBIR1 association for these specific receptors. These examples are consistent with analysis of RLP42 where the Gx_3Gx_3G can be mutated while maintaining a degree of SOBIR1 association [36]. Presence or absence of transmembrane motifs may tune SOBIR1 association levels, but further research is needed.

In conclusion, SOBIR1 exhibits complementary characteristics that allow interaction with 1) the acidic domain E of LRR-RLPs through the SOBIR1 LRRs which have an opposite charge, and 2) helix-helix interaction with their transmembrane domains [36,56]. Although depending on the LRR-RLP, the absence of one out of two motifs, does not necessarily abolish SOBIR1 complex formation (Gust and Felix 2014, Albert 2019).

9. Cytoplasmic tail domain G

Chimeric and truncated LRR-RLPs suggest a variable role for the cytoplasmic tail (G domain) depending on the specific receptor. Although all identified INR homologs encode a cytoplasmic tail of only 10 AA, swapping the INR cytoplasmic tail to the extended version of INR-like 1 did not affect ligand response [24]. Additionally, RLP42 chimeras replacing the terminal LRR, transmembrane helix, and cytoplasmic tail with the those of a non-responsive paralog was still responsive to the pg9(At) elicitor [23]. Additionally, the complete deletion of the intracellular 17-amino-acid tail of RLP23 reduced but did not abolish receptor function [36]. This is consistent with an auxiliary rather than essential function for at least the RLP23 cytoplasmic tail. In contrast, the

deletion of the cytoplasmic tail of either Ve1 and RXEG1 abolishes the function of either receptor [21,37]. Additionally, replacing the Ve1 cytoplasmic tail with the Ve2 version also abolished HR when co-expressed with its ligand Ve1 in tobacco leaves [37]. Hence, the cytoplasmic tail seems to be at least required for Ve1 and RXEG1 mediated resistance.

10. Future Directions

Since their original discovery as resistance genes mediating strain-specific disease resistance, the variety of characterized LRR-RLP functions has expanded to include many different developmental and PAMP sensing roles. Structure-function analyses have begun to delineate specific recognition and signaling roles for the domains described above. An exciting recent development has been cryoEM structural data which indicate a direct role for a conserved ID motif in co-receptor recruitment. We propose that the two classes of K-x_5-Y and Y-x_8-KG motif LRR-RLPs serve as dedicated immune sensors, distinguishing them from development-related LRR-RLPs. The motifs should serve as useful features for characterization of LRR-RLP repertoires. Additional structures are needed, which may require creative strategies to express and purify often recalcitrant LRR-RLP proteins [63]. Finally, we note that LRR-RLPs are often overlooked in resistance gene prioritization pipelines, and are often identified only through targeted reverse genetic analyses using a defined PAMP elicitor [64]. A renewed focus on LRR-RLP discovery and utilization will therefore complement strategies to stack NLR- and RK-based resistance genes [25,65,66]. The LRR-RLP structure provides a versatile platform for recognition of diverse plants, and additional insights into recognition and signaling mechanisms will enable their use for durable disease and pest resistance.

11. Box 1 - SOBIR1 association with LRR-RLPs and potential functional roles

A critical signaling factor required for immunity-related LRR-RLP function is the LRR-RK SUPPRESSOR OF BIR1-1 (SOBIR1). Originally identified as a suppressor of cell death in the *Arabidopsis thaliana bir1* autoimmune background [67], SOBIR1 is now known to constitutively associate with immunity-related RLPs. Previous reviews have summarized the role of SOBIR1 in mediating RLP function [56,68,69]. More recently, constitutive interaction of SOBIR1 with immunity-related LRR-RLPs have been described in species within the Solanaceae, Brassicaceae, and Fabaceae [17,41,47,60,70–73]. SOBIR1 also forms complexes with subfamily VII receptor-like cytoplasmic kinases [53,54], but mechanisms of signal transduction are not well understood.

Importantly, SOBIR1 was secondarily discovered as EVERSHED, a negative regulator of floral organ abscission within a regulatory pathway involving the receptors HAESA and HAESA-LIKE2 [74–76]. SOBIR1 also interacts with the development-related LRR-RLPs CLV2 and TMM [42]. CLV2 canonically acts in meristem maintenance, but its interaction with SOBIR1 may explain alternative immune-related phenotypes in knockout lines [77]. CLV2 is better understood in terms of its interaction with the LRR-RLK CORYNE (CRN) which is known to interact in heteromers with

CLV2 to regulate meristem cell identity [78,79]. Thus, both immunity- and development-related LRR-RLPs seem to signal through subfamily XI-2 LRR-RLKs.

SOBIR1 and CRN are homologs and are the only members of the LRR XI-2 subfamily of Pelle RLKs, based on previous classifications of the large RLK/Pelle family of plant protein kinases [80,81]. However, the origin of separate SOBIR1 and CRN genes by potential gene duplication is not clear. To resolve the evolutionary history of SOBIR1 and CRN in land plants, we constructed a phylogeny of SOBIR1 and CRN homologs after mining the genomes of 34 representative species (Fig. 3). The number of identified SOBIR1 homologs was broadly consistent with LRR-RLK XI-2 identification in the iTAK database of RKs [82].

This phylogenetic analysis revealed two well supported clades within the LRR XI-2 subfamily of Pelle RLKs: one containing characterized SOBIR1 homologs and one containing the characterized Arabidopsis thaliana CRN. We therefore named these clades “SOBIR1-like” and “CRN-like,”. Several independent duplications of SOBIR1 have occurred, for example *Citrus sinensis* (sweet orange) encodes 10 clustered SOBIR1-like genes. We also observed several duplications in the Asterids. Importantly SOBIR1 homologs were found in representatives of the major groups of land plants, including bryophytes such as *Physcomitrella patens, Anthoceros punctatus*, and *Marchantia polymorpha*, as well as lycophytes and ferns, consistent with previous studies into SOBIR1 evolution [57]. In contrast, CRN-like homologs were not identified in these non-seed plant lineages. This suggests broad conservation of SOBIR1-like homologs throughout land plant evolution, but roles of early branching homologs in either immunity or development are not clear.

SOBIR1 homologs share well-characterized motifs critical for RLP association (Supplementary Fig. 1). A basic domain located extracellularly near the transmembrane domain of SOBIR1 is hypothesized to interact with a complementary acidic domain found in RLPs [63]. However EFR-Cf-9 chimeras, suggest that this basic juxtamembrane domain may not be necessary [58]. Within the single pass transmembrane helix of SOBIR1, a Gx_3G motif has been shown to be necessary for interactions with RLP23 [57], mirroring the role of this motif in LRR-RLPs.These two motifs are conserved in SOBIR1-like homologs to different extents (Supplementary Fig. 1). The basic domain is generally conserved throughout land plants, but varies among lineages. In monocots, this domain is histidine-rich, whereas in gymnosperms and eudicots, this domain is lysine-rich. Considering the difference in pH sensitivity, this may help specify monocots and eudicot-specific optima if the juxtamembrane domain is involved in LRR-RLP interactions.

The kinase domain was much more strongly conserved throughout land plants (Supplementary Fig. 2). Specific residues implied in substrate specificity and phosphorylation [83], such as Thr-529 were conserved in all land plant SOBIR1 homologs, but not in CRN. This is consistent with the observation that CRN is kinase inactive. However, phosphorylation sites are variable which may affect signal decoding in different species.

Investigations into cross-species compatibility between SOBIR1 and RLP homologs remain limited. *Arabidopsis thaliana* SOBIR1 (AtSOBIR1) has been shown to interact with *Vigna unguiculata* INR and *Vigna unguiculata* SOBIR1, and transient expression of either AtSOBIR1

and *Phaseolus vulgaris* SOBIR1 equivalently complement transiently expressed *Phaseolus vulgaris* INR loss-of-function in *Nicotiana benthamiana sobir1* loss of function lines [17,84]. Conversely, coexpression of AtSOBIR1 with RLP30 boosts immune signaling output compared to co-expression with SlSOBIR1 of NtSOBIR1, suggesting possible species-specific SOBIR1 and RLP compatibility [85].

Acknowledgements

We thank other members of the Steinbrenner lab for helpful conversations about receptor biology. The work was supported by grant IOS-2139986 to ADS.

Figure Legends

Fig. 1 – Comparative analysis of the NT loopouts and IDs of immunity-related LRR-RLPs with an identified PAMP ligand. a) The AlphaFold2 (AF2) predicted and Cryo-EM (RXEG1) structures of the B- and C-domain of six immunity related LRR-RLPs. If present, the NT loopouts are colored in red. The island domains (IDs) are highlighted in blue. For all LRR-RLPs besides RXEG1, a protein structure prediction was created using the ColabFold platform (v1.3.0; Mirdita et al., 2022). The AF2 input sequence alignment was generated through MMseqs2 using the unpaired+paired mode without using templates (Jumper et al., 2021; Mirdita et al., 2019; Mirdita et al., 2017; Mitchell et al., 2020). Three recycles were run for each of the five created models by AF2-ptm. The five resulting models were ranked based on the pLDDT score, and the highest scoring one was visualized using PyMOL (v2.5.2). **b)** Partial visualization of a mafft e-insi alignment of the complete coding sequences. A full version of the alignment can be found in Supplementary file 2. The transition between the B- and C-domain is indicated with a black vertical line. The main two positions of NT loopouts are highlighted with a red frame. Cys residues within the NT loopouts are highlighted in gray. **c)** ID region of the same alignment. The ID is highlighted with a blue frame. Sequence motifs are added below the IDs of both the K-x_5-Y and Y-x_8-KG type (see section 6) and were generated using WebLogo.

Fig. 2 – LRR-RLP IDs encode K-x_5-Y and Y-x_8-KG motifs with strong phylogenetic structure in four plant species. RLP42 protein sequence was used as TBLASTN query against predicted coding sequences of a) *Arabidopsis thaliana* (Araport11), b) *Populus trichocarpa (*v.4.1, Phytozome genome ID 533), c) *Nicotiana benthamiana* (v.1.01), or d) Oryza sativa (v.7.0, Phytozome genome ID 323). For each species, the top 500 hits were identified and the full protein sequence of any gene hits was collected. Full LRR-RLP protein sequences were aligned using Clustal Omega and a phylogenetic tree was generated using FastTree. A single clade of genes containing no kinase domain and containing either K-x_5-Y or Y-x_8-KG motif IDs was identified for each tree. Numbers indicate bootstrap support separating K-x_5-Y and Y-x_8-KG clades. ID sequences from each species-specific clade (shown as triangles) were manually re-aligned to remove gaps, and sequence logos were generated using WebLogo server (https://weblogo.berkeley.edu/logo.cgi). Phylogenies, alignments, and outgroups can be found in Suppl. File 4.

Fig. 3 – Phylogenetic analysis of subfamily XI-2 RKs reveals deep conservation of separate SOBIR1 and CRN clades. To collect putative homologous sequences, the genomes of representative taxa were searched by TBLASTN using the amino acid sequence of both the SOBIR1 and CRN kinase domains. We used genomes hosted on the Phytozome database, except for a few taxa. *Anthoceros* genomes were accessed from a database hosted by the University of Zurich (https://www.hornworts.uzh.ch/en/hornwort-genomes.html). *Salvinia* and *Azolla* genomes were accessed from the Fernbase database (https://fernbase.org/). The *Ginkgo biloba* genome was accessed from the GinkgoDB database (https://ginkgo.zju.edu.cn/genome/). Finally, the *Helianthus annuus* XQR2.0-SUNRISE genome was accessed from EnsemblPlants (https://plants.ensembl.org/Helianthus_annuus/Info/Annotation/). Results were sorted by e-value and the top ten hits from each of the CRN and SOBIR1 kinase domain TBLASTN searches were collected, for a total of twenty sequences per genome. Duplicate sequences were removed prior to alignment and phylogenetic analysis.

Protein sequences were aligned using the online version of MAFFT (Katoh et al., 2019; Kuraku et al., 2013). A maximum likelihood tree was subsequently generated on the CIPRES web portal using RAXML-HPC2 on XSEDE (v8.2.12) (Miller et al., 2010; Stamatakis, 2014) with the automatic protein model assignment algorithm using maximum likelihood criterion and 100 bootstrap replicates. The resulting phylogeny was rooted and visualized using MEGA11. The subtree representing the Clade XI LRR RLKS was exported and edited in Inkscape.

Supplemental Table Legends

Supplementary table 1:
Characterized LRR-RLPs involved in plant immunity. Excel file which contains the names, annotation, plant species, ligands, pathogen species, certain domain sequences, motif in ID and structure source. NT loopouts and IDs are predicted based on standardized cut-offs in an alignment and visual validation using alphafold2 structure predictions and the cryoEM structure of RXEG1 (Fig. 1, Supplementary file 2, Supplementary file 3). Transmembrane domains were predicted using DeepTMHMM [86].

Supplementary file 1
Fasta file of the amino acid (AA) sequences of LRR-RLPs involved in plant immunity.

Supplementary file 2
Fasta file of the generated alignment of the LRR-RLPs involved in plant immunity. MAFFT 7 alignment using the E-INS-i strategy (Katoh et al., 2002).

Supplementary file 3
Zip file containing all the pdb files of the LRR-RLP structures discussed in Supplementary table 1.

Supplementary file 4

Zip file containing newick trees and fasta alignments of K-x_5-Y and Y-x_8-KG type RLPs from Fig. 2. Alignments and phylogenies include outgroups.

References

[1] Ngou BPM, Ding P, Jones JDG. Thirty years of resistance: Zig-zag through the plant immune system. Plant Cell 2022;34:1447–78. https://doi.org/10.1093/plcell/koac041.

[2] Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant–pathogen interactions. Nat Rev Genet 2010;11:539–48. https://doi.org/10.1038/nrg2812.

[3] Monteiro F, Nishimura MT. Structural, Functional, and Genomic Diversity of Plant NLR Proteins: An Evolved Resource for Rational Engineering of Plant Immunity. Annu Rev Phytopathol 2018;56:243–67. https://doi.org/10.1146/annurev-phyto-080417-045817.

[4] Tamborski J, Krasileva KV. Evolution of Plant NLRs: From Natural History to Precise Modifications. Annu Rev Plant Biol 2020;71:355–78. https://doi.org/10.1146/annurev-arplant-081519-035901.

[5] Albert I, Hua C, Nürnberger T, Pruitt RN, Zhang L. Surface Sensor Systems in Plant Immunity. Plant Physiol 2020;182:1582–96. https://doi.org/10.1104/pp.19.01299.

[6] Boutrot F, Zipfel C. Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. Annu Rev Phytopathol 2017;55:257–86. https://doi.org/10.1146/annurev-phyto-080614-120106.

[7] Shiu SH, Bleecker AB. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Plant Physiol 2003;132:530–43. https://doi.org/10.1104/pp.103.021964.

[8] Smakowska-Luzan E, Mott GA, Parys K, Stegmann M, Howton TC, Layeghifard M, et al. An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. Nature 2018;553:342–6. https://doi.org/10.1038/nature25184.

[9] Chinchilla D, Zipfel C, Robatzek S, Kemmerling B, Nürnberger T, Jones JDG, et al. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. Nature 2007;448:497–500. https://doi.org/10.1038/nature05999.

[10] Hohmann U, Lau K, Hothorn M. The Structural Basis of Ligand Perception and Signal Activation by Receptor Kinases. Annu Rev Plant Biol 2017;68:109–37. https://doi.org/10.1146/annurev-arplant-042916-040957.

[11] Dievart A, Gottin C, Périn C, Ranwez V, Chantret N. Origin and Diversity of Plant Receptor-Like Kinases. Annu Rev Plant Biol 2020;71:131–56. https://doi.org/10.1146/annurev-arplant-073019-025927.

[12] Fritz-Laylin LK, Krishnamurthy N, Tör M, Sjölander KV, Jones JDG. Phylogenomic analysis of the receptor-like proteins of rice and Arabidopsis. Plant Physiol 2005;138:611–23. https://doi.org/10.1104/pp.104.054452.

[13] Wang G, Ellendorff U, Kemp B, Mansfield JW, Forsyth A, Mitchell K, et al. A genome-wide functional investigation into the roles of receptor-like proteins in Arabidopsis. Plant Physiol 2008;147:503–17. https://doi.org/10.1104/pp.108.119487.

[14] Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JD. Isolation of the tomato Cf-9 gene for resistance to Cladosporium fulvum by transposon tagging. Science 1994;266:789–93. https://doi.org/10.1126/science.7973631.

[15] Staskawicz BJ, Ausubel FM, Baker BJ, Ellis JG, Jones JD. Molecular genetics of plant disease resistance. Science 1995;268:661–7. https://doi.org/10.1126/science.7732374.

[16] Albert I, Böhm H, Albert M, Feiler CE, Imkampe J, Wallmeroth N, et al. An RLP23-SOBIR1-BAK1 complex mediates NLP-triggered immunity. Nat Plants 2015;1:15140. https://doi.org/10.1038/nplants.2015.140.

[17] Steinbrenner AD, Muñoz-Amatriaín M, Chaparro AF, Aguilar-Venegas JM, Lo S, Okuda S, et al. A receptor-like protein mediates plant immune responses to herbivore-associated molecular patterns. Proc Natl Acad Sci U S A 2020. https://doi.org/10.1073/pnas.2018415117.

[18] Hegenauer V, Slaby P, Körner M, Bruckmüller J-A, Burggraf R, Albert I, et al. The tomato receptor CuRe1 senses a cell wall protein to identify Cuscuta as a pathogen. Nat Commun 2020;11:5299. https://doi.org/10.1038/s41467-020-19147-4.

[19] Nie J, Zhou W, Liu J, Tan N, Zhou J-M, Huang L. A receptor-like protein from Nicotiana benthamiana mediates VmE02 PAMP-triggered immunity. New Phytol 2020. https://doi.org/10.1111/nph.16995.

[20] Zhang L, Kars I, Essenstam B, Liebrand TWH, Wagemakers L, Elberse J, et al. Fungal endopolygalacturonases are recognized as microbe-associated molecular patterns by the arabidopsis receptor-like protein RESPONSIVENESS TO BOTRYTIS POLYGALACTURONASES1. Plant Physiol 2014;164:352–64. https://doi.org/10.1104/pp.113.230698.

[21] Sun Y, Wang Y, Zhang X, Chen Z, Xia Y, Wang L, et al. Plant receptor-like protein activation by a microbial glycoside hydrolase. Nature 2022;610:335–42. https://doi.org/10.1038/s41586-022-05214-x.

[22] Fan L, Fröhlich K, Melzer E, Pruitt RN, Albert I, Zhang L, et al. Genotyping-by-sequencing-based identification of Arabidopsis pattern recognition receptor RLP32 recognizing proteobacterial translation initiation factor IF1. Nat Commun 2022;13:1294. https://doi.org/10.1038/s41467-022-28887-4.

[23] Zhang L, Hua C, Pruitt RN, Qin S, Wang L, Albert I, et al. Distinct immune sensor systems for fungal endopolygalacturonases in closely related Brassicaceae. Nat Plants 2021;7:1254–63. https://doi.org/10.1038/s41477-021-00982-2.
[24] Snoeck S, Abramson BW, Garcia AGK, Egan AN, Michael TP, Steinbrenner AD. Evolutionary gain and loss of a plant pattern-recognition receptor for HAMP recognition. Elife 2022. https://doi.org/10.7554/eLife.81050.
[25] van Esse HP, Reuber TL, van der Does D. Genetic modification to improve disease resistance in crops. New Phytol 2020;225:70–86. https://doi.org/10.1111/nph.15967.
[26] Dangl JL, Horvath DM, Staskawicz BJ. Pivoting the plant immune system from dissection to deployment. Science 2013;341:746–51. https://doi.org/10.1126/science.1236011.
[27] Steidele C, Stam R. Multi-omics approach highlights differences between functional RLP classes in Arabidopsis thaliana 2020:2020.08.07.240911. https://doi.org/10.1101/2020.08.07.240911.
[28] van der Hoorn RAL, Wulff BBH, Rivas S, Durrant MC, van der Ploeg A, de Wit PJGM, et al. Structure-function analysis of cf-9, a receptor-like protein with extracytoplasmic leucine-rich repeats. Plant Cell 2005;17:1000–15. https://doi.org/10.1105/tpc.104.028118.
[29] Jones DA, Jones JDG. The Role of Leucine-Rich Repeat Proteins in Plant Defences. In: Andrews JH, Tommerup IC, Callow JA, editors. Advances in Botanical Research, vol. 24, Academic Press; 1997, p. 89–167. https://doi.org/10.1016/S0065-2296(08)60072-5.
[30] Martin R, Qi T, Zhang H, Liu F, King M, Toth C, et al. Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. Science 2020;370. https://doi.org/10.1126/science.abd9993.
[31] Yoon S-I, Kurnasov O, Natarajan V, Hong M, Gudkov AV, Osterman AL, et al. Structural basis of TLR5-flagellin recognition and signaling. Science 2012;335:859–64. https://doi.org/10.1126/science.1215584.
[32] Matsushima N, Miyashita H. Leucine-Rich Repeat (LRR) Domains Containing Intervening Motifs in Plants. Biomolecules 2012;2:288–311. https://doi.org/10.3390/biom2020288.
[33] Wulff BB, Thomas CM, Smoker M, Grant M, Jones JD. Domain swapping and gene shuffling identify sequences required for induction of an Avr-dependent hypersensitive response by the tomato Cf-4 and Cf-9 proteins. Plant Cell 2001;13:255–72. https://doi.org/10.1105/tpc.13.2.255.
[34] Van der Hoorn RA, Roth R, De Wit PJ. Identification of distinct specificity determinants in resistance protein Cf-4 allows construction of a Cf-9 mutant that confers recognition of avirulence protein Avr4. Plant Cell 2001;13:273–85. https://doi.org/10.1105/tpc.13.2.273.
[35] Wulff BBH, Heese A, Tomlinson-Buhot L, Jones DA, de la Peña M, Jones JDG. The major specificity-determining amino acids of the tomato Cf-9 disease resistance protein are at hypervariable solvent-exposed positions in the central leucine-rich repeats. Mol Plant Microbe Interact 2009;22:1203–13. https://doi.org/10.1094/MPMI-22-10-1203.
[36] Albert I, Zhang L, Bemm H, Nürnberger T. Structure-Function Analysis of Immune Receptor AtRLP23 with its Ligand nlp20 and Coreceptors AtSOBIR1 and AtBAK1. Mol Plant Microbe Interact 2019:MPMI09180263R. https://doi.org/10.1094/MPMI-09-18-0263-R.
[37] Fradin EF, Zhang Z, Rovenich H, Song Y, Liebrand TWH, Masini L, et al. Functional analysis of the tomato immune receptor Ve1 through domain swaps with its non-functional homolog Ve2. PLoS One 2014;9:e88208. https://doi.org/10.1371/journal.pone.0088208.
[38] Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K, Jones JD. The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 1996;84:451–9. https://doi.org/10.1016/s0092-8674(00)81290-8.
[39] Dixon MS, Hatzixanthis K, Jones DA, Harrison K, Jones JD. The tomato Cf-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. Plant Cell 1998;10:1915–25.
[40] Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, et al. Characterization of the tomato Cf-4 gene for resistance to Cladosporium fulvum identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. Plant Cell 1997;9:2209–24. https://doi.org/10.1105/tpc.9.12.2209.
[41] van der Burgh AM, Postma J, Robatzek S, Joosten MHAJ. Kinase activity of SOBIR1 and BAK1 is required for immune signalling. Mol Plant Pathol 2019;20:410–22. https://doi.org/10.1111/mpp.12767.
[42] Liebrand TWH, van den Berg GCM, Zhang Z, Smit P, Cordewener JHG, America AHP, et al. Receptor-like kinase SOBIR1/EVR interacts with receptor-like proteins in plant immunity against fungal infection. Proc Natl Acad Sci U S A 2013;110:10010–5. https://doi.org/10.1073/pnas.1220015110.
[43] Fradin EF, Zhang Z, Juarez Ayala JC, Castroverde CDM, Nazar RN, Robb J, et al. Genetic dissection of Verticillium wilt resistance mediated by tomato Ve1. Plant Physiol 2009;150:320–32. https://doi.org/10.1104/pp.109.136762.
[44] Zhang W, Fraiture M, Kolb D, Löffelhardt B, Desaki Y, Boutrot FFG, et al. Arabidopsis receptor-like protein30 and receptor-like kinase suppressor of BIR1-1/EVERSHED mediate innate immunity to necrotrophic fungi. Plant Cell 2013;25:4227–41. https://doi.org/10.1105/tpc.113.117010.
[45] Bar M, Sharfman M, Ron M, Avni A. BAK1 is required for the attenuation of ethylene-inducing xylanase (Eix)-induced defense responses by the decoy receptor LeEix1. Plant J 2010;63:791–800. https://doi.org/10.1111/j.1365-313X.2010.04282.x.
[46] Postma J, Liebrand TWH, Bi G, Evrard A, Bye RR, Mbengue M, et al. Avr4 promotes Cf-4 receptor-like protein

association with the BAK1/SERK3 receptor-like kinase to initiate receptor endocytosis and plant immunity. New Phytol 2016;210:627–42. https://doi.org/10.1111/nph.13802.
[47] Du J, Verzaux E, Chaparro-Garcia A, Bijsterbosch G, Keizer LCP, Zhou J, et al. Elicitin recognition confers enhanced resistance to Phytophthora infestans in potato. Nat Plants 2015;1:15034. https://doi.org/10.1038/nplants.2015.34.
[48] Jamieson PA, Shan L, He P. Plant cell surface molecular cypher: Receptor-like proteins and their roles in immunity and development. Plant Sci 2018;274:242–51. https://doi.org/10.1016/j.plantsci.2018.05.030.
[49] Steinbrenner AD. The evolving landscape of cell surface pattern recognition across plant immune networks. Curr Opin Plant Biol 2020;56:135–46. https://doi.org/10.1016/j.pbi.2020.05.001.
[50] Wang G, Long Y, Thomma BPHJ, de Wit PJGM, Angenent GC, Fiers M. Functional analyses of the CLAVATA2-like proteins and their domains that contribute to CLAVATA2 specificity. Plant Physiol 2010;152:320–31. https://doi.org/10.1104/pp.109.148197.
[51] Lin G, Zhang L, Han Z, Yang X, Liu W, Li E, et al. A receptor-like protein acts as a specificity switch for the regulation of stomatal development. Genes Dev 2017;31:927–38. https://doi.org/10.1101/gad.297580.117.
[52] Cammarata J, Scanlon MJ. A functionally informed evolutionary framework for the study of LRR-RLKs during stem cell maintenance. J Plant Res 2020;133:331–42. https://doi.org/10.1007/s10265-020-01197-w.
[53] Pruitt RN, Locci F, Wanke F, Zhang L, Saile SC, Joe A, et al. The EDS1–PAD4–ADR1 node mediates Arabidopsis pattern-triggered immunity. Nature 2021:1–5. https://doi.org/10.1038/s41586-021-03829-0.
[54] Tian H, Wu Z, Chen S, Ao K, Huang W, Yaghmaiean H, et al. Activation of TIR signalling boosts pattern-triggered immunity. Nature 2021;598:500–3. https://doi.org/10.1038/s41586-021-03987-1.
[55] Kourelis J, Contreras MP, Harant A, Pai H, Lüdke D, Adachi H, et al. The helper NLR immune protein NRC3 mediates the hypersensitive cell death caused by the cell-surface receptor Cf-4. PLoS Genet 2022;18:e1010414. https://doi.org/10.1371/journal.pgen.1010414.
[56] Gust AA, Felix G. Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. Curr Opin Plant Biol 2014;21:104–11. https://doi.org/10.1016/j.pbi.2014.07.007.
[57] Bi G, Liebrand TWH, Bye RR, Postma J, van der Burgh AM, Robatzek S, et al. SOBIR1 requires the GxxxG dimerization motif in its transmembrane domain to form constitutive complexes with receptor-like proteins. Mol Plant Pathol 2016;17:96–107. https://doi.org/10.1111/mpp.12266.
[58] Wu J, Reca I-B, Spinelli F, Lironi D, De Lorenzo G, Poltronieri P, et al. An EFR-Cf-9 chimera confers enhanced resistance to bacterial pathogens by SOBIR1- and BAK1-dependent recognition of elf18. Mol Plant Pathol 2019;20:751–64. https://doi.org/10.1111/mpp.12789.
[59] Russ WP, Engelman DM. The GxxxG motif: a framework for transmembrane helix-helix association. J Mol Biol 2000;296:911–9. https://doi.org/10.1006/jmbi.1999.3489.
[60] Wang Y, Xu Y, Sun Y, Wang H, Qi J, Wan B, et al. Leucine-rich repeat receptor-like gene screen reveals that Nicotiana RXEG1 regulates glycoside hydrolase 12 MAMP detection. Nat Commun 2018;9:594. https://doi.org/10.1038/s41467-018-03010-8.
[61] Jehle AK, Lipschis M, Albert M, Fallahzadeh-Mamaghani V, Fürst U, Mueller K, et al. The receptor-like protein ReMAX of Arabidopsis detects the microbe-associated molecular pattern eMax from Xanthomonas. Plant Cell 2013;25:2330–40. https://doi.org/10.1105/tpc.113.110833.
[62] Catanzariti A-M, Do HTT, Bru P, de Sain M, Thatcher LF, Rep M, et al. The tomato I gene for Fusarium wilt resistance encodes an atypical leucine-rich repeat receptor-like protein whose function is nevertheless dependent on SOBIR1 and SERK3/BAK1. Plant J 2017;89:1195–209. https://doi.org/10.1111/tpj.13458.
[63] Hohmann U, Hothorn M. Crystal structure of the leucine-rich repeat ectodomain of the plant immune receptor kinase SOBIR1. Acta Crystallogr D Struct Biol 2019;75:488–97. https://doi.org/10.1107/S2059798319005291.
[64] Schultink A, Steinbrenner AD. A playbook for developing disease-resistant crops through immune receptor identification and transfer. Curr Opin Plant Biol 2021;62:102089. https://doi.org/10.1016/j.pbi.2021.102089.
[65] Lacombe S, Rougon-Cardoso A, Sherwood E, Peeters N, Dahlbeck D, van Esse HP, et al. Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nat Biotechnol 2010;28:365–9. https://doi.org/10.1038/nbt.1613.
[66] Luo M, Xie L, Chakraborty S, Wang A, Matny O, Jugovich M, et al. A five-transgene cassette confers broad-spectrum resistance to a fungal rust pathogen in wheat. Nat Biotechnol 2021. https://doi.org/10.1038/s41587-020-00770-x.
[67] Gao M, Wang X, Wang D, Xu F, Ding X, Zhang Z, et al. Regulation of cell death and innate immunity by two receptor-like kinases in Arabidopsis. Cell Host Microbe 2009;6:34–44. https://doi.org/10.1016/j.chom.2009.05.019.
[68] Zhang Z, Thomma BPHJ. Structure-function aspects of extracellular leucine-rich repeat-containing cell surface receptors in plants. J Integr Plant Biol 2013;55:1212–23. https://doi.org/10.1111/jipb.12080.
[69] Liebrand TWH, van den Burg HA, Joosten MHAJ. Two for all: receptor-associated kinases SOBIR1 and BAK1. Trends Plant Sci 2014;19:123–32. https://doi.org/10.1016/j.tplants.2013.10.003.
[70] Catanzariti A-M, Lim GTT, Jones DA. The tomato I-3 gene: a novel gene for resistance to Fusarium wilt disease. New Phytol 2015;207:106–18. https://doi.org/10.1111/nph.13348.
[71] Hegenauer V, Fürst U, Kaiser B, Smoker M, Zipfel C, Felix G, et al. Detection of the plant parasite Cuscuta

reflexa by a tomato cell surface receptor. Science 2016;353:478–81. https://doi.org/10.1126/science.aaf3919.
[72] Domazakis E, Wouters D, Visser RGF, Kamoun S, Joosten MHAJ, Vleeshouwers VGAA. The ELR-SOBIR1 Complex Functions as a Two-Component Receptor-Like Kinase to Mount Defense Against Phytophthora infestans. Mol Plant Microbe Interact 2018:MPMI09170217R. https://doi.org/10.1094/MPMI-09-17-0217-R.
[73] Ma L, Borhan MH. The receptor-like kinase SOBIR1 interacts with Brassica napus LepR3 and is required for Leptosphaeria maculans AvrLm1-triggered immunity. Front Plant Sci 2015;6:933. https://doi.org/10.3389/fpls.2015.00933.
[74] Leslie ME, Lewis MW, Youn J-Y, Daniels MJ, Liljegren SJ. The EVERSHED receptor-like kinase modulates floral organ shedding in Arabidopsis. Development 2010;137:467–76. https://doi.org/10.1242/dev.041335.
[75] Gubert CM, Liljegren SJ. HAESA and HAESA-LIKE2 activate organ abscission downstream of NEVERSHED and EVERSHED in Arabidopsis flowers. Plant Signal Behav 2014;9:e29115. https://doi.org/10.4161/psb.29115.
[76] Patharkar OR, Walker JC. Floral organ abscission is regulated by a positive feedback loop. Proc Natl Acad Sci U S A 2015;112:2906–11. https://doi.org/10.1073/pnas.1423595112.
[77] Pan L, Lv S, Yang N, Lv Y, Liu Z, Wu J, et al. The Multifunction of CLAVATA2 in Plant Development and Immunity. Front Plant Sci 2016;7:1573. https://doi.org/10.3389/fpls.2016.01573.
[78] Müller R, Bleckmann A, Simon R. The receptor kinase CORYNE of Arabidopsis transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. Plant Cell 2008;20:934–46. https://doi.org/10.1105/tpc.107.057547.
[79] Bleckmann A, Weidtkamp-Peters S, Seidel CAM, Simon R. Stem cell signaling in Arabidopsis requires CRN to localize CLV2 to the plasma membrane. Plant Physiol 2010;152:166–76. https://doi.org/10.1104/pp.109.149930.
[80] Lehti-Shiu MD, Shiu S-H. Diversity, classification and function of the plant protein kinase superfamily. Philos Trans R Soc Lond B Biol Sci 2012;367:2619–39. https://doi.org/10.1098/rstb.2012.0003.
[81] Liu P-L, Du L, Huang Y, Gao S-M, Yu M. Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. BMC Evol Biol 2017;17:47. https://doi.org/10.1186/s12862-017-0891-5.
[82] Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, et al. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Mol Plant 2016;9:1667–70. https://doi.org/10.1016/j.molp.2016.09.014.
[83] Wei X, Wang Y, Zhang S, Gu T, Steinmetz G, Yu H, et al. Structural analysis of receptor-like kinase SOBIR1 reveals mechanisms that regulate its phosphorylation-dependent activation. Plant Communications 2022:100301. https://doi.org/10.1016/j.xplc.2022.100301.
[84] Garcia AGK, Steinbrenner AD. Bringing Plant Immunity to Light: A Genetically Encoded, Bioluminescent Reporter of Pattern Triggered Immunity in Nicotiana benthamiana. Mol Plant Microbe Interact 2022. https://doi.org/10.1094/MPMI-07-22-0160-TA.
[85] Yang Y, Steidele CE, Löffelhardt B, Kolb D, Leisen T, Zhang W, et al. Convergent evolution of plant pattern recognition receptors sensing cysteine-rich patterns from three microbial kingdoms. bioRxiv 2022:2022.10.06.511083. https://doi.org/10.1101/2022.10.06.511083.
[86] Hallgren J, Tsirigos KD, Pedersen MD, Armenteros JJA, Marcatili P, Nielsen H, et al. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv 2022:2022.04.08.487609. https://doi.org/10.1101/2022.04.08.487609.

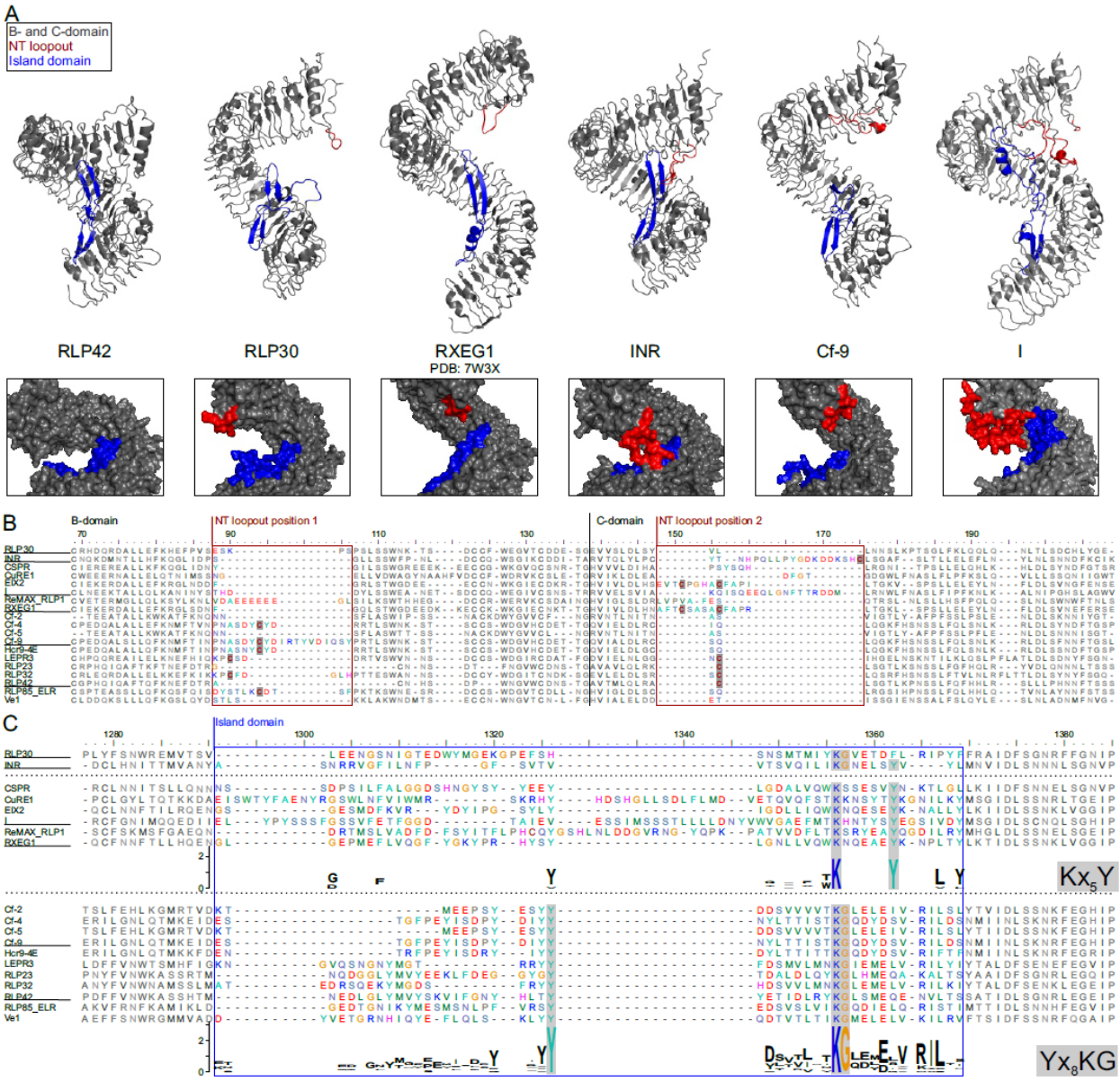

Fig. 1 – Comparative analysis of the NT loopouts and IDs of immunity-related LRR-RLPs with an identified PAMP ligand. a) The AlphaFold2 (AF2) predicted and Cryo-EM (RXEG1) structures of the B- and C-domain of six immunity related LRR-RLPs. If present, the NT loopouts are colored in red. The island domains (IDs) are highlighted in blue. For all LRR-RLPs besides RXEG1, a protein structure prediction was created using the ColabFold platform (v1.3.0; Mirdita et al., 2022). The AF2 input sequence alignment was generated through MMseqs2 using the unpaired+paired mode without using templates (Jumper et al., 2021; Mirdita et al., 2019; Mirdita et al., 2017; Mitchell et al., 2020). Three recycles were run for each of the five created models by AF2-ptm. The five resulting models were ranked based on the pLDDT score, and the highest scoring one was visualized using PyMOL (v2.5.2). **b)** Partial visualization of a mafft e-insi alignment of the complete coding sequences. A full version of the alignment can be found in Supplementary file 2. The transition between the B- and C-domain is indicated with a black vertical line. The main two positions of NT loopouts are highlighted with a red frame.

Cys residues within the NT loopouts are highlighted in gray. **c)** ID region of the same alignment. The ID is highlighted with a blue frame. Sequence motifs are added below the IDs of both the K-x_5-Y and Y-x_8-KG type (see section 6) and were generated using WebLogo.

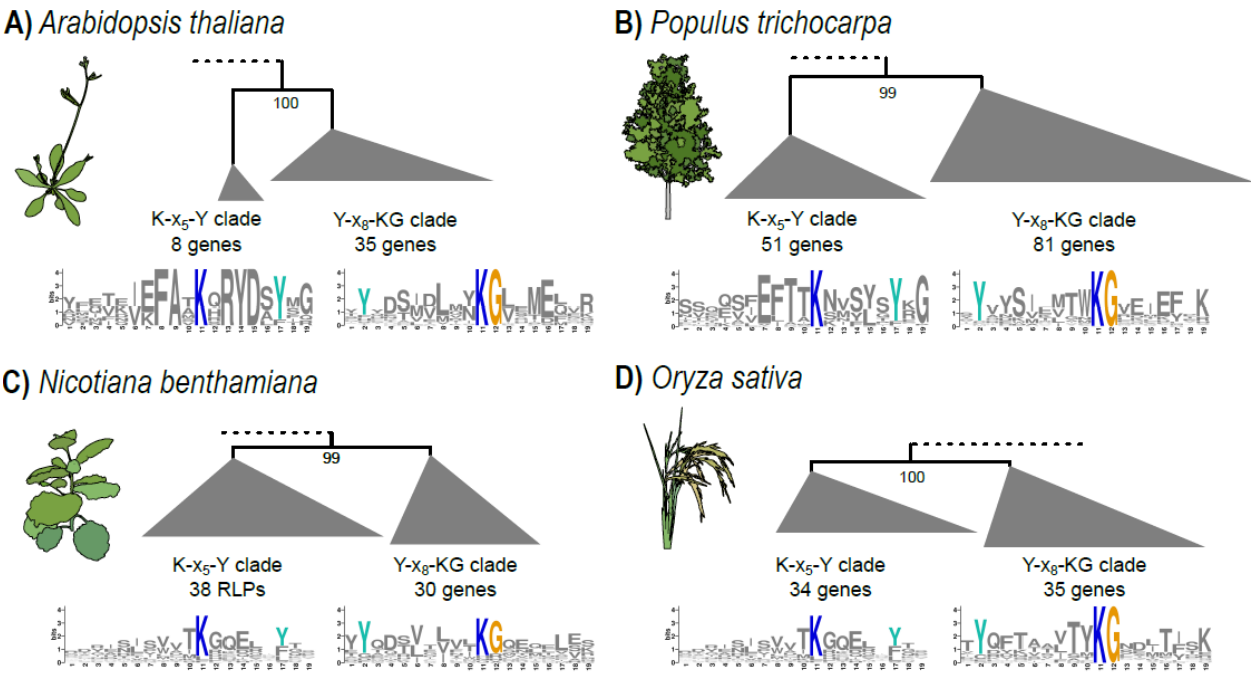

Fig. 2 – LRR-RLP IDs encode K-x$_5$-Y and Y-x$_8$-KG motifs with strong phylogenetic structure in four plant species. RLP42 protein sequence was used as TBLASTN query against predicted coding sequences of a) *Arabidopsis thaliana* (Araport11), b) *Populus trichocarpa (*v.4.1, Phytozome genome ID 533), c) *Nicotiana benthamiana* (v.1.01), or d) Oryza sativa (v.7.0, Phytozome genome ID 323). For each species, the top 500 hits were identified and the full protein sequence of any gene hits was collected. Full LRR-RLP protein sequences were aligned using Clustal Omega and a phylogenetic tree was generated using FastTree. A single clade of genes containing no kinase domain and containing either K-x$_5$-Y or Y-x$_8$-KG motif IDs was identified for each tree. Numbers indicate bootstrap support separating K-x$_5$-Y and Y-x$_8$-KG clades. ID sequences from each species-specific clade (shown as triangles) were manually re-aligned to remove gaps, and sequence logos were generated using WebLogo server (https://weblogo.berkeley.edu/logo.cgi). Phylogenies, alignments, and outgroups can be found in Suppl. File 4.

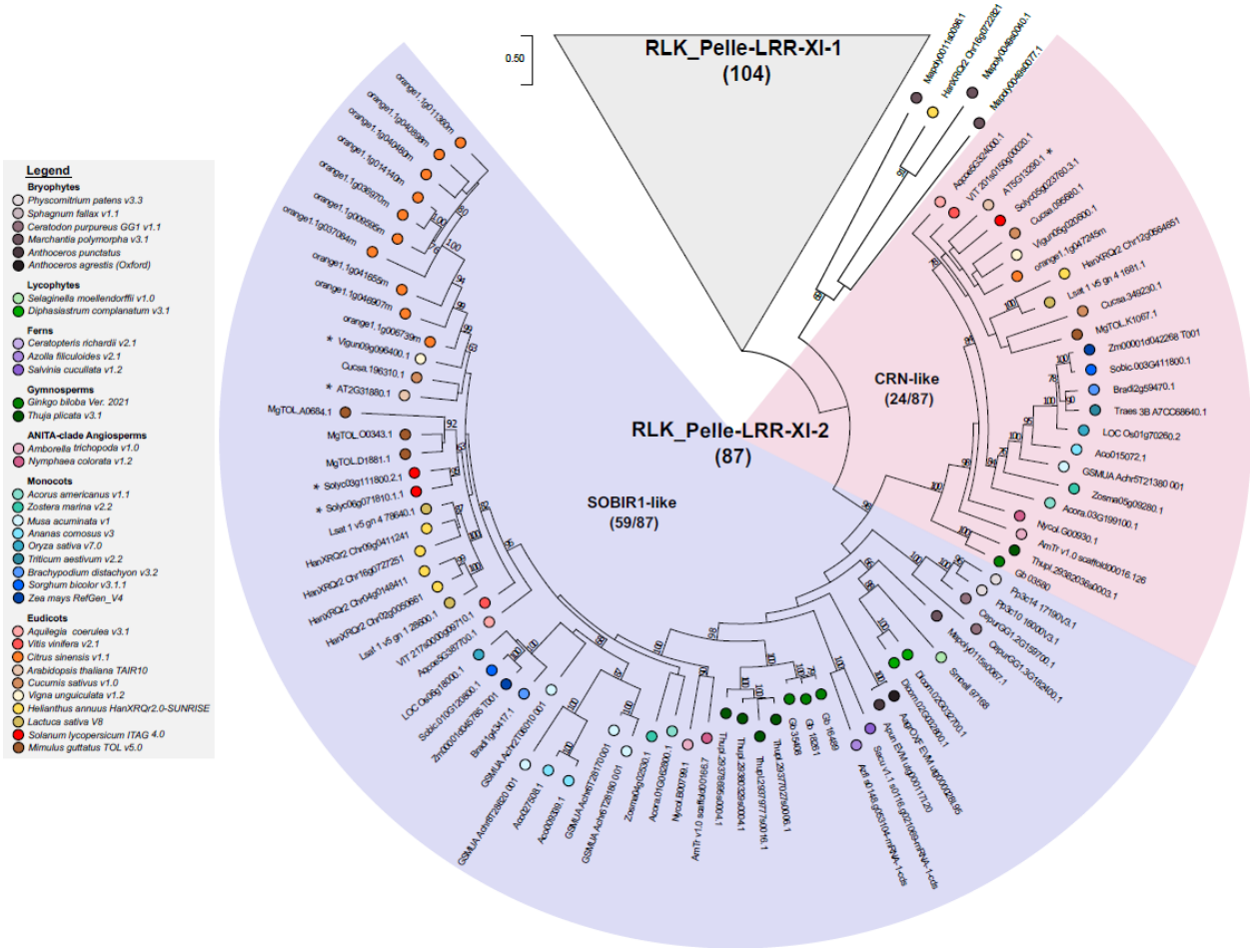

Fig. 3 – Phylogenetic analysis of subfamily XI-2 RKs reveals deep conservation of separate SOBIR1 and CRN clades. To collect putative homologous sequences, the genomes of representative taxa were searched by TBLASTN using the amino acid sequence of both the SOBIR1 and CRN kinase domains. We used genomes hosted on the Phytozome database, except for a few taxa. *Anthoceros* genomes were accessed from a database hosted by the University of Zurich (https://www.hornworts.uzh.ch/en/hornwort-genomes.html). *Salvinia* and *Azolla* genomes were accessed from the Fernbase database (https://fernbase.org/). The *Ginkgo biloba* genome was accessed from the GinkgoDB database (https://ginkgo.zju.edu.cn/genome/). Finally, the *Helianthus annuus* XQR2.0-SUNRISE genome was accessed from EnsemblPlants (https://plants.ensembl.org/Helianthus_annuus/Info/Annotation/). Results were sorted by e-value and the top ten hits from each of the CRN and SOBIR1 kinase domain TBLASTN searches were collected, for a total of twenty sequences per genome. Duplicate sequences were removed prior to alignment and phylogenetic analysis.