


Differential expansion and retention patterns of LRR-RLK genes across plant evolution

Zachary Kileeg^{1,2} | Aparna Haldar^{1,2} | Hasna Khan^{1,2} | Arooj Qamar¹ | G. Adam Mott^{1,2,3} 

¹Department of Biological Sciences, University of Toronto - Scarborough, Toronto, Canada

²Department of Cell and Systems Biology, University of Toronto, Toronto, Canada

³Centre for the Analysis of Genome Evolution & Function, University of Toronto, Toronto, Canada

Correspondence

G. Adam Mott, Department of Biological Sciences, University of Toronto - Scarborough, Toronto, ON, Canada.

Email: adam.mott@utoronto.ca

Funding information

This work was supported by Natural Sciences and Engineering Research Council of Canada through a Discovery Grant Award to GAM and an Undergraduate Student Research Award to HK.

Abstract

To maximize overall fitness, plants must accurately respond to a host of growth, developmental, and environmental signals throughout their life. Many of these internal and external signals are perceived by the leucine-rich repeat receptor-like kinases, which play roles in regulating growth, development, and immunity. This largest family of receptor kinases in plants can be divided into subfamilies based on the conservation of the kinase domain, which demonstrates that shared evolutionary history often indicates shared molecular function. Here we investigate the evolutionary history of this family across the evolution of 112 plant species. We identify lineage-specific expansions of the malectin-domain containing subfamily LRR subfamily I primarily in the Brassicales and bryophytes. Most other plant lineages instead show a large expansion in LRR subfamily XII, which in *Arabidopsis* is known to contain key receptors in pathogen perception. This striking asymmetric expansion may reveal a dichotomy in the evolutionary history and adaptation strategies employed by plants. A greater understanding of the evolutionary pressures and adaptation strategies acting on members of this receptor family offers a way to improve functional predictions for orphan receptors and simplify the identification of novel stress-related receptors.

1 | INTRODUCTION

In order to survive, plants must appropriately respond to numerous extracellular signals. These signals guide both development and immunity in the plant. To aid in discriminating between these signals, plants have evolved a large suite of cell surface receptors. Many of these receptors are receptor-like kinases (RLKs), with the largest group in *Arabidopsis* (*Arabidopsis thaliana*) being the leucine-rich repeat receptor-like kinase family (LRR-RLK) (Lehti-Shiu et al., 2009; Shiu & Bleecker, 2001).

The LRR-RLKs have three protein domains: an N-terminal extracellular domain (ECD) to receive signals and mediate the

physical interactions that lead to the formation of receptor complexes, a single transmembrane domain (TM), and a C-terminal intracellular kinase domain (KD). Previous analyses have divided the LRR-RLKs into numerous subfamilies based on an inferred phylogeny from the alignment of their KDs (Liu et al., 2017; Shiu & Bleecker, 2001). The LRR-RLKs are named for the presence of numerous leucine-rich repeats (LRR) within the ECD. In addition to these canonical repeats, the ECDs of LRR-RLK-I (LRR-I), LRR-RLK-VIII.1 (LRR-VIII.1) and LRR-RLK-VIII.2 (LRR-VIII.2) subfamily members contain malectin, or malectin-like, domains (MD/MLDs). MDs show sequence similarity to the malectins, which are small ER-resident disaccharide-binding proteins in animal cells (Schallus et al., 2008, 2010), while the MLD is a plant-specific domain composed of two tandem MDs.

Zachary Kileeg and Aparna Haldar, contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Plant Direct* published by American Society of Plant Biologists and the Society for Experimental Biology and John Wiley & Sons Ltd.

As they often impact agronomically important traits, the LRR-RLKs have long been seen as attractive targets for crop improvement (Diévar & Clark, 2004; Lemmon et al., 2018; Rodríguez-Leal et al., 2017). In spite of this interest, only a minority of LRR-RLKs have defined biological functions. The majority of LRR-RLKs that have been characterized to date are broadly involved in either immunity or growth and development (Dufayard et al., 2017; Fischer et al., 2016; Tang et al., 2010). An impediment to functional characterization is the large number of closely related genes with the potential for high rates of functional redundancy (Nimchuk et al., 2015; Rodríguez-Leal et al., 2019).

The expansion and diversification of LRR-RLK repertoires within plant lineages have resulted from both whole genome duplications and smaller segmental or tandem duplications (Panchy et al., 2016). Whole genome duplication has occurred many times over the course of angiosperm evolution (Renny-Byfield & Wendel, 2014; Soltis et al., 2009), while tandem and segmental duplications are also common (Rizzon et al., 2006). Interestingly, rates of LRR-RLK duplication and retention appear correlated to their biological function (Lehti-Shiu et al., 2009; Ngou et al., 2022). Specifically, subfamilies involved in stress responses are greatly expanded by tandem duplication, while developmental LRR-RLKs lack this signature (Shiu et al., 2004; Tang et al., 2010). The genes from expanded subfamilies also show evidence of increased positive selection when compared to LRR-RLKs from nonexpanded subfamilies (Dufayard et al., 2017; Fischer et al., 2016; Tang et al., 2010). This selection is observed predominantly in the ECD, which is responsible for signal recognition (Dufayard et al., 2017; Fischer et al., 2016; Tang et al., 2010).

In this work, we examine the expansion of the LRR-RLK family across plant evolution. We show a specific expansion of the LRR-I subfamily within the Brassicales that is driven mainly by tandem duplication. Most other plant lineages show expansion of the LRR-RLK-XII (LRR-XII) subfamily, which contains the best-studied LRR-RLKs involved in pathogen perception and response. Interestingly, almost all plant species show a preferential expansion of either the LRR-I or LRR-XII family. We then further examine the diversity of LRR-I protein sequences, demonstrating two distinct domain architectures and the presence of a well-conserved cleavage motif across plant lineages.

2 | METHODS

2.1 | Plant genomes and LRR-RLK identification

Representative protein sequences were retrieved from Phytozome version 13 (Goodstein et al., 2012) for the 112 plant and algae species available as of May 2020. These 112 species were assembled into an approximate species tree using the NCBI taxonomy browser (Schoch et al., 2020) with red algae (*Porphyra umbilicalis*, species code: Pumbi) as the outgroup. Each species was assigned a five-letter code for ease of identification (gene codes and species names can be found in Table S1) and divided into taxonomic groups for downstream analysis. Species were assigned to their respective orders within angiosperms where appropriate, with mosses and liverworts assigned to the bryophytes,

and green algae to Chlorophyta. Orders containing two or fewer species were grouped together into a single group termed 'Other'.

To identify LRR-RLK genes from the list of representative protein sequences we first identified proteins that contained an LRR domain using predict-phytoLRR (Chen, 2021), a protein kinase domain (PF00069, PF07714), or an NB-ARC (PF00931) domain using HMMSearch (Eddy, 2011) (v3.3). We included all proteins containing both an LRR and kinase domain, but lacking an NB-ARC domain in the collection of LRR-RLKs. The kinase domains of all LRR-RLK sequences were aligned using MAFFT (Katoh & Standley, 2013) (v.7.453), and an unrooted phylogeny was inferred using fasttree (Price et al., 2010) (v2.1.11) with default settings. Using this tree we assigned the LRR-RLKs to subfamilies using Arabidopsis LRR-RLK annotations as a guide (Wang et al., 2019). In cases where existing Arabidopsis subfamilies resolved into two or more clades, the smaller group of the subfamily was split and designated into a new division of that subfamily. We then calculated the proportion of LRR-RLKs belonging to each subfamily in each species.

2.2 | Analysis of tandem duplication

We identified any LRR-RLKs that arose via tandem duplication by manual inspection. We defined tandem duplicates as LRR-RLKs belonging to the same subfamily on the same chromosome separated by no more than two intervening genes with at least 50% similarity between their entire protein-coding sequences.

2.3 | LRR-I and LRR-XII phylogenies

We inferred separate phylogenetic trees for LRR-I and LRR-XII subfamilies using sequences from all 112 species. The kinase domains of the LRR-RLK protein sequences were aligned using default settings in MAFFT (Katoh & Standley, 2013) (v.7.453) and maximum-likelihood phylogenies were constructed using the JTT + F + R10 substitution model based on model finder (Kalyaanamoorthy et al., 2017) in IQ-TREE (Nguyen et al., 2015) with 1,000 bootstrap replicates (v1.6.12). LRR-I trees were rooted using a variety of human and Arabidopsis kinase sequences (Sakamoto et al., 2012). Subclades were assigned based on a bottom-up approach where sequences from new branches were merged with the growing subclade if they shared an average of 60% sequence similarity and a bootstrap support of at least 90%. This was done until the next branch to be added contained sequence similarity <60%.

2.4 | Ancestral state reconstruction

We used ancestral state reconstruction to infer the maximum-likelihood number of LRR-I or LRR-XII subfamily members at each branch of the NCBI species tree using fastAnc from the R package phytools (Revell, 2012) (v0.7-70). The branch-specific expansion rate for each subfamily was calculated by dividing the number of LRR-



RLKs at a given node by the number of LRR-RLKs of the same subfamily in the most recent ancestral branch.

2.5 | Motif analysis

We input the LRR-I protein sequences from all 112 species into MEME from memesuite (Bailey et al., 2015) (v5.4.1) and considered the most significant results that contained at least 5 amino acids in the motif. The protein sequences were compared against the UniProtKB protein database to examine motif prevalence using ScanProsite (de Castro et al., 2006) (v20.0). Occurrences of exact motif matches were separated by subfamily and counted.

2.6 | Analysis of domain architecture

LRR-I protein sequences for each subclade were analyzed with HMMScan in HMMER (Eddy, 2011) (v.3.3) using the pfam-a HMM set (Mistry et al., 2021). All significant domain hits (E-value <1) were output and counted. Transmembrane regions were predicted using TMHMM (Krogh et al., 2001) (v2.0), and signal peptides were predicted using SignalP (Almagro Armenteros et al., 2019) (v5.0) for representative sequences. All predicted domains were compared and consensus domains were plotted using a representative sequence. Disulfide bonds were inferred from the predicted protein structure of the Arabidopsis FRK1 protein (AF-O64483-F1). Unique domains were plotted on the most common architecture in the general area where the domain appears.

2.7 | Protein structure mapping

The predicted protein structure of FRK1 (AF-O64483-F1) was obtained from AlphaFold (Jumper et al., 2021) (v2.0), and the transmembrane region and cytosolic region removed from the model in PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). The locations of motifs and other sequence features from MEME or HMMScan were mapped to the extracellular region in PyMOL.

2.8 | Tissue expression analysis

Tissue expression analysis for 42 Arabidopsis LRR-I genes was performed using e-Northern (Toufighi et al., 2005) using default settings.

3 | RESULTS

3.1 | LRR-RLKs display subfamily-specific expansion patterns across 112 plant and algal species

Based on their known domain structure, we identified LRR-RLK genes as those that encode proteins with an LRR domain and a kinase

domain, but lacking an NB-ARC domain. This definition was used to identify all putative LRR-RLK genes in 112 plant and algal genomes. The protein sequences of the kinase domains from all putative LRR-RLK genes for each species were aligned and subfamily designations were determined based on similarity to known Arabidopsis genes. The number of LRR-RLK genes present in each subfamily for each species was used to identify subfamily-specific rates of gene expansion and retention. We looked at the proportion of genes in each subfamily and found that the majority of subfamilies have consistent proportions across all plant species studied (Figure 1a, Table S2). Six subfamilies deviate from the others, showing considerable variation between plant species and containing large numbers of expanded genes (Figure 1a, subfamilies LRR-I, LRR-III, LRR-VIII.1, LRR-VIII.2, LRR-XI.1b, LRR-XI.1c, and LRR-XII). These subfamilies have the highest proportion of LRR-RLKs and display the greatest variability in gene proportion across the sampled species, both expected characteristics of genes involved in stress responses.

We next sought to determine the mechanism of subfamily expansion, as tandem duplication and subsequent retention are also correlated with stress response genes (Hanada et al., 2008; Qiao et al., 2019). In Arabidopsis, the expansions observed in subfamilies LRR-I, LRR-VIII.1, LRR-VIII.2, and LRR-XII are associated with high rates of tandem duplication (>50% of subfamily genes), suggestive of potential roles in stress responses (Figure 1b).

3.2 | The expansions of subfamilies LRR-I and LRR-XII are largely negatively correlated

Having identified several subfamilies with evidence of expansion through tandem duplication, we chose to focus on subfamily LRR-I. LRR-I was chosen because it is the largest subfamily within Arabidopsis, has extensive tandem duplication, and has members known to function in biotic stress responses (Asai et al., 2002; Yeh et al., 2016; Yuan et al., 2018). As a comparator, we chose subfamily LRR-XII, as it has a similar expansion pattern to LRR-I, shows high tandem duplication, and is known to contain receptors involved in the biotic stress response, including FLS2 (Chinchilla et al., 2006; Mott et al., 2016; Zipfel et al., 2006). To explore the lineage-based expansion dynamic of these subfamilies, we grouped the sequences from LRR-I and LRR-XII into their respective orders and inferred unrooted maximum-likelihood phylogenies of both (Figure S1). This revealed an asymmetric expansion pattern, where Brassicales has a clear and particularly broad expansion in LRR-I but not in LRR-XII, whereas Poales shows the opposite trend (Figure S1). To avoid any sampling effects arising from input species number, we explored this asymmetry by comparing the proportions of LRR-I and LRR-XII receptors across orders (Figure 2a, Figure S2). As expected, Poales has significantly more LRR-XII genes than LRR-I (Figure 2a, Figure S2). Similarly, the Malvales, Rosales, and Sapindales orders all have significantly more LRR-XII receptors than LRR-I (Figure 2a, Figure S2). In contrast, the opposite trend is seen in Brassicales and bryophytes, in which we identify significantly more LRR-I than LRR-XII (Figure 2a, Figure S2). The order Brassicales and the bryophytes are notable for their relative dearth of

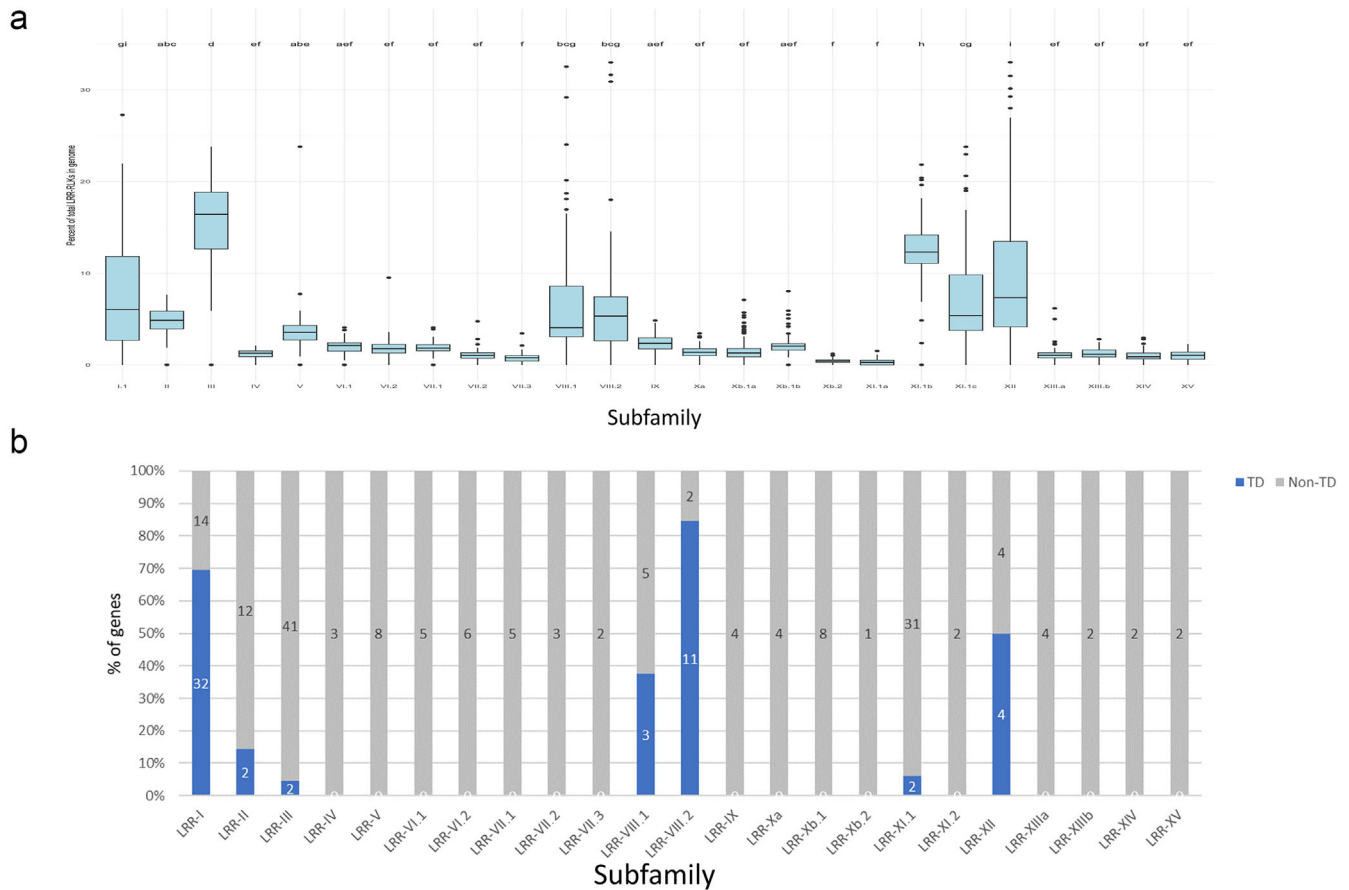


FIGURE 1 Expansion, duplication, and counts of LRR-RLKs across 112 plant and algae species. (a) The LRR-RLK subfamily members have undergone asymmetric expansion. Subfamily expansion is represented as the percentage of LRR-RLKs belonging to each subfamily out of the total number of LRR-RLKs in each genome. Letters represent the significance group of each subfamily, as determined by Tukey's HSD test at $\alpha = .05$. Subfamilies with shared letters are statistically similar to one another, while subfamilies with different letters are significantly different from one another. Species with less than 10 total LRR-RLKs were omitted. Subfamilies LRR-I, LRR-III, LRR-VIII, LRR-XI, and LRR-XII have a higher proportion of total LRR-RLKs compared to the other subfamilies, demonstrating increased diversification. (b) LRR-RLKs have undergone tandem duplication. Blue bars represent tandem duplicated genes, while gray bars represent non-tandem duplicated genes. Tandem duplicates are those in the same subclade separated on a chromosome by no more than two intervening genes.

subfamily LRR-XII members compared to other plant lineages, suggesting a possible negative correlation between expansions in these two subfamilies.

We next tested whether this observed expansion asymmetry was broadly true in plants by comparing the ratio of LRR-I and LRR-XII gene count in each plant species (Figure 2b – external nodes). The majority of species show a large expansion in subfamily LRR-XII without a similar expansion in LRR-I (shown in blue). Fewer independent plant groups show the inverse pattern, which is like that observed in *Arabidopsis* and the bryophytes (red). We observe fewer lineages with similar expansions in these two subfamilies (gray). To estimate at what point during plant evolution these patterns emerged, we used ancestral state reconstruction to infer branch-specific expansion for each species (Figure 2b – internal nodes, Figure S3). We found that the high number of LRR-I genes present in modern-day Brassicales, especially in the Brassicaceae family may be the result of two expansion events. The first is in the ancestor of Brassicales and the second is in the ancestor of Brassicaceae, followed by local duplication events

(Figure 2b, Figure S3). Concomitant with this expansion of LRR-I, there is a reduction in the number of LRR-XII genes in the Brassicales. In contrast, numerous non-Brassicales angiosperms show expansion in LRR-XII rather than LRR-I (Figure 2). In particular, the order Poales shows the largest expansion of subfamily LRR-XII among any of the plant groups studied (Figure 2b, Figure S3).

3.3 | There are 25 well-supported subclades within the LRR-I subfamily

Based on the preferential expansion of LRR-I in Brassicales and bryophytes, we chose to investigate the evolutionary relationship of LRR-I sequences. Of the 112 species of plant and algae we investigated, 101 contain at least one LRR-I gene. We used the sequences from these 101 species to reconstruct the phylogeny of LRR-I. We identified 25 subclades within LRR-I supported by high sequence similarity and widespread bootstrap support (Figure 3, Table S3). Subclades

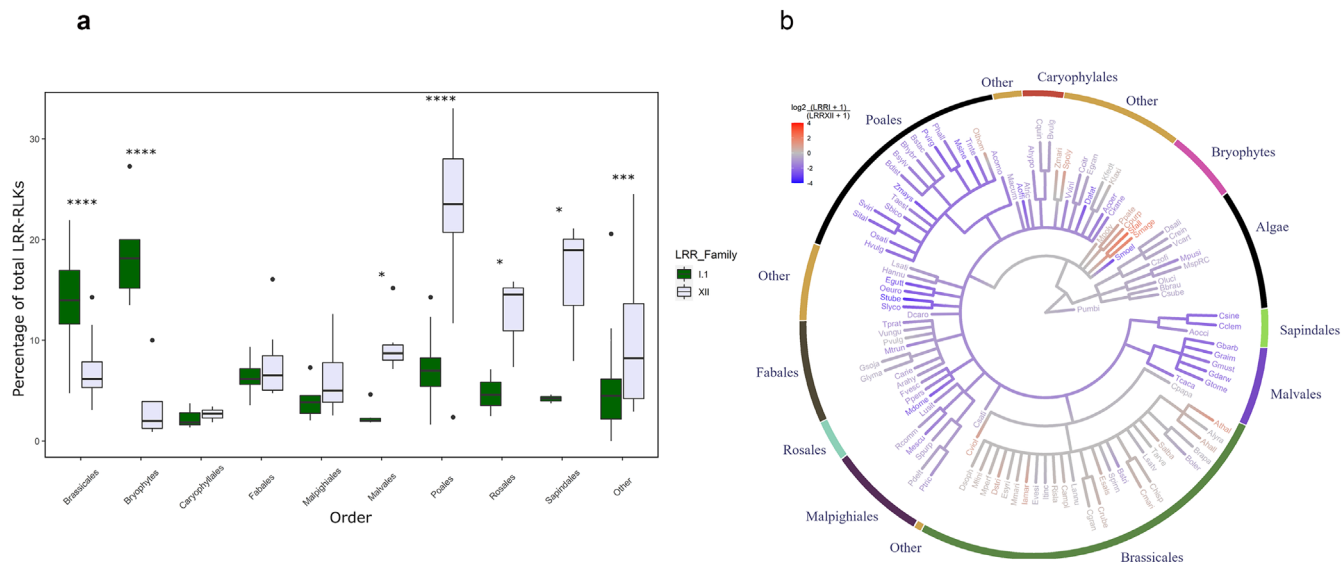


FIGURE 2 Comparison of LRR-I and LRR-XII receptor counts across different plant orders. (a) Percentage of total LRR-RLKs found in subfamilies LRR-I and LRR-XII in each given order. Outliers ($> Q3 + 1.5 \times IQR$) are plotted as dots. Significance values represent an ANOVA followed by Tukey's HSD. * = $p < .05$, ** = $p < .01$, *** = $p < .001$, **** = $p < .0001$. (b) LRR-RLK subfamilies have different ratios of expansion across evolutionary time. Colors indicate the log of the ratio of LRR-I to LRR-XII genes at each node. Numbers at terminal nodes represent the ratio for those species and internal nodes represent the ratio for the common ancestor at that branch. Red algae (*P. umbilicalis*) is used as the species outgroup.

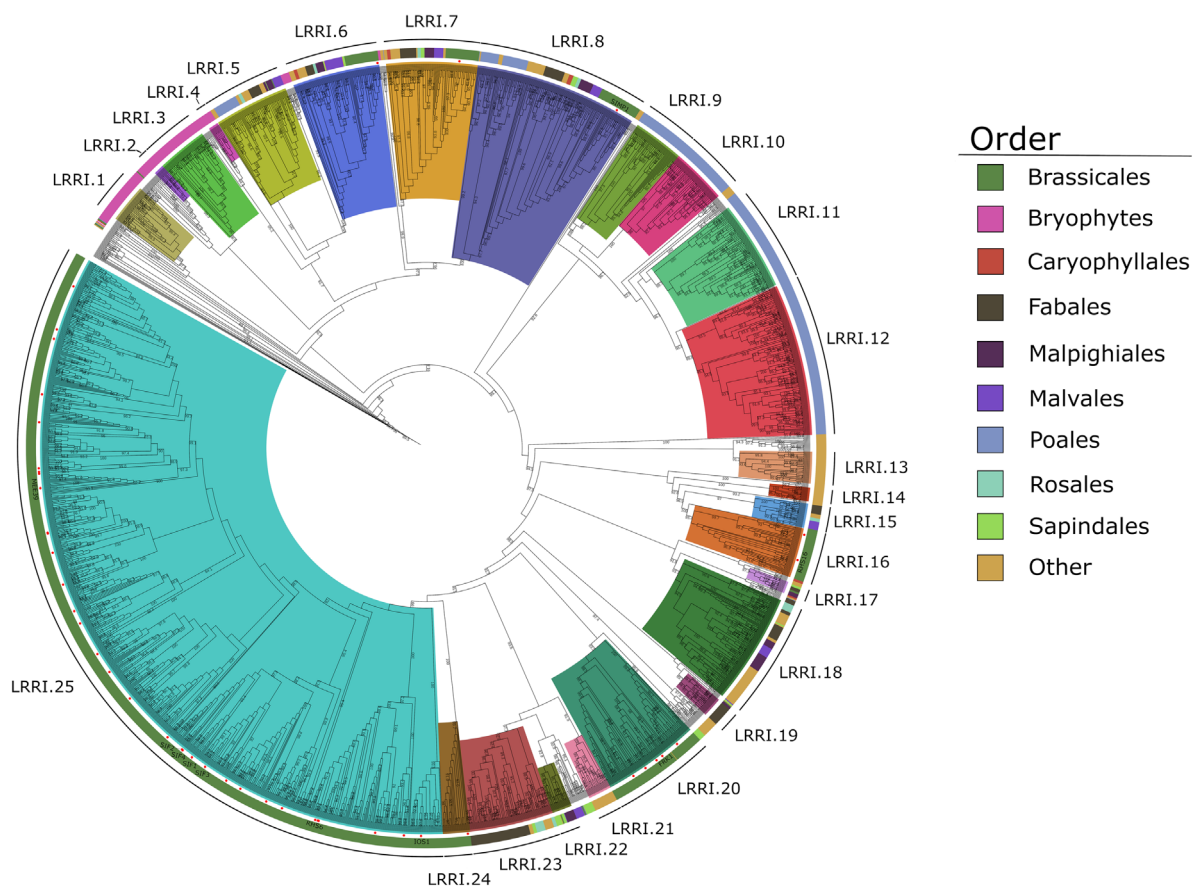


FIGURE 3 Maximum-likelihood phylogeny of LRR-I inferred from 112 species of plants and algae based on kinase domain sequences. Outer ring denotes the order in each sequence falls into. Inner ring colors denote the subclade to which each sequence belongs. Bootstrap values represent 1,000 re-samplings. Subclades were considered as such if sequences shared at least 60% sequence similarity to one another and the subclade contained bootstrap support $>90\%$. Phylogeny was rooted using a variety of Arabidopsis and human kinases. Arabidopsis sequences are marked with red dots, and characterized members are noted.

LRR-I.1 through LRR-I.4 exclusively contain bryophyte sequences, while the remaining bryophyte sequences do not cluster in any well-supported subclades. Subclades LRR-I.5 through LRR-I.8, LRR-I.15, LRR-I.17, LRR-I.18, LRR-I.22, and LRR-I.23 contain sequences from a diverse range of plant species but do not contain any bryophyte sequences (Figure 3). Notably, subclades LRR-I.1 through LRR-I.4, LRR-I.9 through LRR-I.14, LRR-I.16, LRR-I.19 through LRR-I.21, LRR-I.24, and LRR-I.25 are lineage-specific, with >95% of the sequences derived from a single order of plants (Figure 3). Among these, Brassicales and Poales have the largest and second-largest expansions, respectively (Figures 2 and 3). There also appears to be species-specific expansion, denoted by subclades LRR-I.4 and LRR-I.13 through LRR-I.15 although this may be a consequence of limited lineage sampling (Figure 3).

3.4 | Novel motifs and domain architectures

After observing high levels of species and lineage-specific expansion and retention of LRR-I receptor repertoires, we hypothesized this may be evidence of diversification and neofunctionalization within this subfamily. We determined that LRR-I receptors primarily display one of two architectures: the first has an ECD containing a malectin-like domain and a variable number of leucine-rich repeats (2–4) followed by a transmembrane domain and a cytosolic kinase domain, while the second type is either missing the MLD and has a longer LRR domain or contains a truncated MLD (Figure 4a, Figure S4C, Table S3). The second architecture is found almost exclusively in the bryophyte-containing subclade LRR-I.1 (Figure 3). The angiosperm LRR-I proteins

almost universally contain an MLD domain and display the first architecture, indicating a clear division between the bryophytes and flowering plants (Figure 4). Uncommon domains also were found in several regions including the N-terminal region of the MLD, between the repeated malectin domains within the MLD, within gaps between and overlapping with LRR domains, spanning the transmembrane/juxtamembrane region, and within the C-terminal region of the protein (Figure 4b, Figure S4, Table S3).

We next investigated whether there were novel protein motifs found specifically within the ECDs of these proteins. In doing so, we found numerous highly conserved motifs within the extracellular region (Figure S5). The most common non-LRR motif was the WxGDPCxP motif (GDPC, Figure 4b, Table S3). Over 87% of all LRR-I receptors contain this motif located between the well-annotated malectin-like and LRR domains (Figure 4, Figures S6 and S7). The majority of LRR-I sequences contain both the GDPC and MLD. Notably, however, the GDPC motif is mostly absent from receptors in LRR-I.1, LRR-I.6, and LRR-I.23, although LRR-I.6 and LRR-I.23 receptors do contain an MLD (Figure S6).

4 | DISCUSSION

Despite their well-documented roles in growth and immunity, the majority of plant LRR-RLKs still have no known biological function. Assigning functions to individual receptors is time-consuming and complicated by the large number of genes and the potential for functional redundancy. We hope that the use of evolutionary and phylogenetic signatures to identify candidate LRR-RLKs that are likely to play

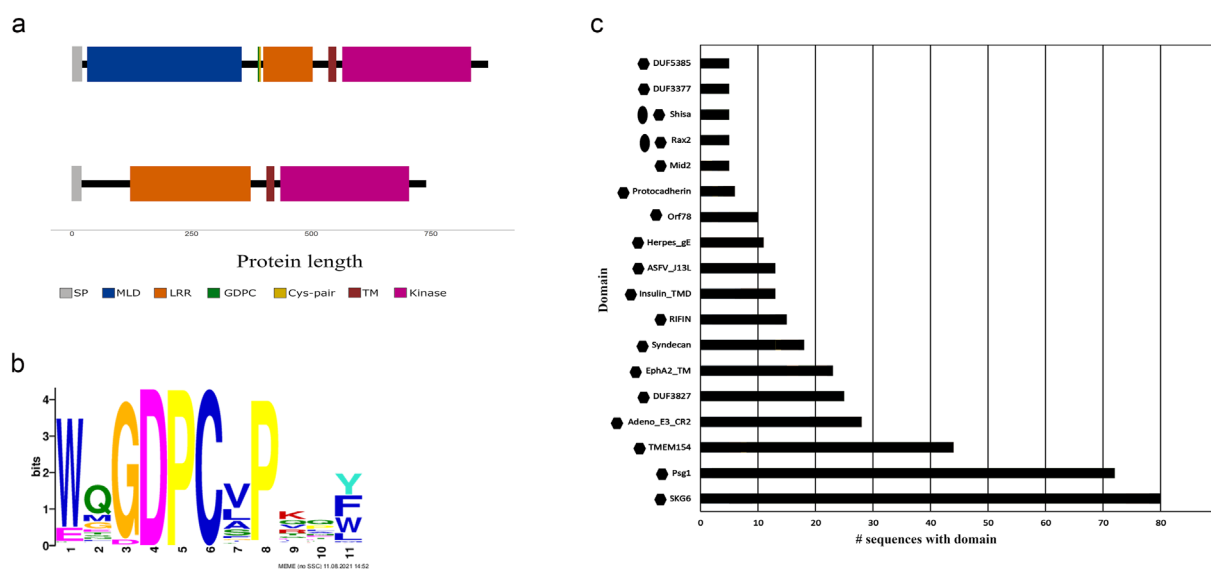


FIGURE 4 Domain architectures of LRR-I members across 112 species. (a) General schematic of the two commonly found architectures across 112 species. MLD = malectin-like domain, LRRs = leucine-rich repeats, TM = transmembrane domain, GDPC = a cleavage motif, Cys-pair = predicted disulfide bond, and SP = signal peptide. The two architectures are based on Arabidopsis FRK1 (top) and *P. patens* receptor Ppate_Pp3c16_23380V3.1.p (bottom). (b) Sequence logo of the WxGDPCxP motif found in the majority of LRR-I sequences. (c) Uncommon domains found within at least five sequences. Hexagons represent those found within the transmembrane domain, and ovals represent those found within the LRR domain.



roles in adaptation to stress may streamline this process. We therefore investigated the LRR-RLK repertoires of 112 plant and algal species to identify the subfamilies most likely to contain stress response receptors, which we hypothesize to be those with high levels of gene expansion, particularly via tandem duplication (Qiao et al., 2019).

In accordance with previous reports (Fischer et al., 2016; Ngou et al., 2022), our analysis showed varying rates of gene expansion and retention both between LRR-RLK gene subfamilies within a species, and in a single LRR-RLK subfamily between different species. While most subfamilies showed low levels of expansion across plant evolution and little variability between species, there were several subfamilies that showed increased gene number, variability in gene content across species, and potential asymmetric expansion, which generally confirm the findings of previous analyses (Fischer et al., 2016). These features were especially prominent in subfamilies LRR-I, LRR-III, LRR-VIII.1, LRR-VIII.2, LRR-XI, and LRR-XII (Figure 1a). These families display both high levels of expansion and between-species variability, and interestingly they all contain at least one member known to be involved in stress responses (Aryal et al., 2023; Chinchilla et al., 2006; Le et al., 2014; Tang et al., 2015; Uemura et al., 2020; Yeh et al., 2016). The inclusion of additional species in this study did identify expansions, particularly in the LRR-VIII.1 and LRR-VIII.2 subfamilies, that were not noted in previous work. While large increases in LRR-VIII.1 numbers were observed in a variety of species including two members of both the Myrtaceae and Crassulaceae families, the effect on LRR-VIII.2 was driven almost exclusively by bryophyte genomes with increased gene number.

We hypothesized that subfamilies that act broadly in stress responses would also show high levels of tandem duplication and subsequent gene retention. Duplication and retention provide genetic material for neofunctionalization and signal diversification to broaden the ability to detect biotic stress in particular (Hanada et al., 2008). While we observed evidence of tandem duplication in seven subfamilies in Arabidopsis, only three had more than one cluster of duplicated genes (Figure 1b). Subfamilies LRR-I, LRR-VIII.2, and LRR-XII are all greatly expanded via tandem duplication, with at least half of the total gene content found in duplicated clusters.

The only other subfamily to contain more than two tandem duplicated genes is LRR-VIII.1. It is interesting to note that LRR-I, LRR-VIII.1, and LRR-VIII.2 are the only three subfamilies that contain either a malectin or malectin-like domain and they have all been expanded via multiple tandem duplication clusters. Malectin domains are named after malectin, a conserved di-saccharide binding protein found in animal cells (Schallus et al., 2008, 2010). The ability to bind carbohydrates, combined with the expansion of the motif within plant genomes, led to speculation that receptor kinases containing these domains may act as cell wall sensors (Yang et al., 2021). Subsequent work has shown that several of these receptors play roles in immunity signaling (Chan et al., 2020; Hok et al., 2011; Robatzek & Somssich, 2002; Yeh et al., 2015, 2016). These findings would support the hypothesis that in Arabidopsis and other Brassicales, the expansion and retention of LRR-I genes is associated with their function in biotic stress response.

The expansion of the LRR-I subfamily in Brassicales has also created several lineage-specific subclades. While most members of the subfamily remain unstudied, the few characterized examples show interesting evolutionary patterns. The flg22-induced receptor-like kinase 1 (FRK1) is a commonly used marker of pattern-triggered immunity (Asai et al., 2002) and is found in LRR-I.20. This subclade is dominated by a Brassicales lineage-specific expansion with only one sequence derived from a different order. Likewise, impaired oomycete susceptibility 1 (IOS1) has been shown to play a role in pattern-triggered immunity broadly in addition to its oomycete phenotype (Hok et al., 2011; Yeh et al., 2016), and is also found within a Brassicales lineage-specific subclade (LRR-I.25). Also present within LRR-I.25 is stress-induced factor 2 (SIF2) and the related SIF-family members. SIF2 has been shown to modulate pattern-triggered immunity and bacterial defense (Chan et al., 2020; Yuan et al., 2018). LRR-I.25 also contains maternal effect embryo arrest mutant 39 (MEE39), which when mutated causes embryo development to arrest at the one-cell stage (Pagnussat et al., 2005). The finding that so many receptors involved in the biotic-stress response in Arabidopsis are found within these highly expanded subclades would support the hypothesis that expansion has been driven by pressures on the immune system. The only well-characterized Arabidopsis LRR-I gene found outside of a Brassicales lineage-specific expansion is salt-induced malectin-like domain-containing protein 1 (SIMP1), which plays a role in salt stress tolerance by elevating ER-associated protein degradation of misfolded proteins and is found in LRR-I.8 (He et al., 2021).

It is not clear why this expansion of LRR-I seems to come at the expense of the expansion of the PRR-containing LRR-XII subfamily seen in most other lineages (Figure 2). Interestingly, the observed expansion in LRR-I coincides with several other genetic and life history changes in the Brassicales lineage. Many Brassicaceae are non-hosts to arbuscular mycorrhizae fungi (AM) or associate with only limited fungal species (Cosme et al., 2018). The Brassicaceae family also shows a unique shoot-skewed expression pattern of the immune-related nucleotide-binding leucine-rich repeat receptors (NLRs). In contrast, other plant species preferentially express these resistance genes in the roots (Munch et al., 2018). The coincident loss of AM fungal association and NLR expression in the roots suggests that Brassicaceae may have evolved a unique mechanism for protecting roots from pathogenic attack. Interestingly, most LRR-I genes in Arabidopsis display preferential expression in the root, positioning them perfectly to provide a non-specific immune response against a wide variety of microbes (Figure S8). In contrast, plants that depend upon associations with commensal fungi may require more specific root protection that can effectively identify pathogenic invaders while allowing beneficial associations. If true, this hypothesis would still raise the question of whether the loss of AM-fungal associations allowed the development of a non-specific root immunity, or whether the development of the immune response eliminated the possibility of commensal association.

To further investigate the possibility of LRR-I receptors gaining novel function, we also looked for the presence of other protein

motifs or domains. The incorporation of such novel features can lead to functional diversity, a phenomenon recently shown for the intracellular NLR family of receptors (Van de Weyer et al., 2019). The majority of the newly identified domains span the transmembrane domain and have no known function. Given their localization, they may assist in receptor complex formation or signal propagation, a phenomenon that has been observed in other kinase systems (Arcas et al., 2020; Hubert et al., 2010; Li & Hristova, 2010).

In addition, we also identified the well-conserved GDPC motif located between the MLD and LRR regions in the majority of LRR-I sequences (Figure 4, Figures S6 and S7). This motif has been previously identified in LRR-RLK subfamilies LRR-I, LRR-IV, LRR-V, and LRR-VIII.1 (Dufayard et al., 2017), but we have identified it in a minority of LRR-VIII.2 receptors (210/1989 LRR-VIII.2 receptors, data not shown). This domain has been studied in depth in the case of the LRR-I member symbiosis receptor-like kinase (SYMRK). In *Lotus japonicus*, the GDPC motif acts as a cleavage signal, resulting in the release of the MLD from the receptor (Antolín-Llovera et al., 2014; Kosuta et al., 2011). The release of the MLD allows the remaining receptor, containing only the LRR domains, to associate with a member of a family of Lysin-motif (LYS-M) containing receptors important for chitin recognition (Buendía et al., 2018). Given the conservation of the motif, we hypothesize that members of the LRR-I subfamily share a common regulatory mechanism dependent on cleavage and release of the MLD region.

While over 87% of all LRR-I receptors contain the GDPC motif, it is not found in LRR-I.1, LRR-I.6, or LRR-I.23 (Figure S4). Although lacking the GDPC motif, LRR-I.6 and LRR-I.23 do contain the MLD (Figure S4). In contrast, LRR-I.1 and LRR-I.4 are the only subclades lacking the MLD (Figure S4). However, LRR-I.4 does contain the GDPC motif (Figure S4). These sequences are found exclusively in *Marchantia* and raise the obvious question of what function cleavage at the GDPC motif may possibly serve in this context. The lack of the GDPC motif in LRR-I.1 but the presence in LRR-I.4 suggests the GDPC motif was acquired prior to the addition of the MLD domain to this subfamily of receptors. Lack of the GDPC motif in LRR-I.6 and LRR-I.23, then, may indicate that receptor cleavage is not always essential.

Through this work, we investigated the LRR-RLK family of receptor kinases across 112 species of plants and algae to find evidence of expansion. In our work, we found LRR-RLK subfamilies underwent asymmetric expansion across different lineages, with LRR-I being preferentially expanded in bryophytes and Brassicales. Furthermore, we found that the GDPC cleavage motif is found in the majority of LRR-I receptors, implying the motif is necessary for these receptors' function. The work presented here provides new insights into the expansion and evolution of the LRR-I family. Further analysis of other lineages, including those within the gymnosperms, would give even more information on the evolution and expansion of the LRR-I family.

AUTHOR CONTRIBUTIONS

Study concept and design: Zachary Kileeg, Adam Mott. *Analysis and interpretation of data:* Zachary Kileeg, Aparna Haldar, Hasna Khan,

Arooj Qamar. *Initial draft of the manuscript:* Zachary Kileeg. *Critical revision of the manuscript:* Zachary Kileeg, Aparna Haldar, Hasna Khan, Adam Mott. *Study supervision:* Adam Mott.

CONFLICT OF INTEREST STATEMENT

The Authors did not report any conflict of interest.

PEER REVIEW

The peer review history for this article is available in the Supporting Information for this article.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

G. Adam Mott  <https://orcid.org/0000-0001-9803-103X>

REFERENCES

- Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 37(4), 420–423.
- Antolín-Llovera, M., Ried, M. K., & Parniske, M. (2014). Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE ectodomain promotes complex formation with nod factor receptor 5. *Current Biology*, 24, 422–427. <https://doi.org/10.1016/j.cub.2013.12.053>
- Arcas, A., Wilkinson, D. G., & Nieto, M. Á. (2020). The evolutionary history of Ephs and Ephrins: Toward multicellular organisms. *Molecular Biology and Evolution*, 37, 379–394. <https://doi.org/10.1093/molbev/msz222>
- Aryal, B., Xia, J., Hu, Z., Stumpe, M., Tsering, T., Liu, J., Huynh, J., Fukao, Y., Glöckner, N., Huang, H. Y., Sáncho-Andrés, G., Pakula, K., Ziegler, J., Gorzalka, K., Zwiewka, M., Nodzynski, T., Harter, K., Sánchez-Rodríguez, C., Jasiński, M., ... Geisler, M. M. (2023). An LRR receptor kinase controls ABC transporter substrate preferences during plant growth-defense decisions. *Current Biology*, 33, 2008–2023.e8. <https://doi.org/10.1016/j.cub.2023.04.029>
- Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W.-L. L., Gomez-Gomez, L., Boller, T., Ausubel, F. M., & Sheen, J. (2002). MAP kinase signalling cascade in *Arabidopsis* innate immunity. *Nature*, 415, 977–983. <https://doi.org/10.1038/415977a>
- Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME suite. *Nucleic Acids Research*, 43, W39–W49. <https://doi.org/10.1093/nar/gkv416>
- Buendía, L., Girardin, A., Wang, T., Cottret, L., & Lefebvre, B. (2018). LysM receptor-like kinase and LysM receptor-like protein families: An update on phylogeny and functional characterization. *Front Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.01531>
- Chan, C., Panzeri, D., Okuma, E., Töldsepp, K., Wang, Y. Y., Louh, G. Y., Chin, T. C., Yeh, Y. H., Yeh, H. L., Yekondi, S., Huang, Y. H., Huang, T. Y., Chiou, T. J., Murata, Y., Kollist, H., & Zimmerli, L. (2020). STRESS INDUCED FACTOR 2 regulates *Arabidopsis* stomatal immunity through phosphorylation of the Anion Channel SLAC1. *Plant Cell*, 32, 2216–2236. <https://doi.org/10.1105/tpc.19.00578>
- Chen, T. (2021). Identification and characterization of the LRR repeats in plant LRR-RLKs. *BMC Molecular and Cellular Biology*, 22, 9. <https://doi.org/10.1186/s12860-021-00344-y>
- Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., & Felix, G. (2006). The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines



- the specificity of flagellin perception. *Plant Cell*, 18, 465–476. <https://doi.org/10.1105/tpc.105.036574>
- Cosme, M., Fernández, I., Van der Heijden, M. G. A., & Pieterse, C. M. J. (2018). Non-mycorrhizal plants: The exceptions that prove the rule. *Trends in Plant Science*, 23, 577–587. <https://doi.org/10.1016/j.tplants.2018.04.004>
- de Castro, E., Sigrist, C. J. A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., Bairoch, A., & Hulo, N. (2006). ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research*, 34, W362–W365. <https://doi.org/10.1093/nar/gkl124>
- Diévert, A., & Clark, S. E. (2004). LRR-containing receptors regulating plant development and defense. *Development*, 131, 251–261. <https://doi.org/10.1242/dev.00998>
- Dufayard, J.-F. F., Bettembourg, M., Fischer, I., Droc, G., Guiderdoni, E., Périn, C., Chantret, N., & Diévert, A. (2017). New insights on leucine-rich repeats receptor-like kinase orthologous relationships in angiosperms. *Frontiers in Plant Science*, 08, 381.
- Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, 7, e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>
- Fischer, I., Diévert, A., Droc, G., Dufayard, J.-F., Chantret, N., Diévert, A., Droc, G., Dufayard, J.-F., & Chantret, N. (2016). Evolutionary dynamics of the leucine-rich repeat receptor-like kinase (LRR-RLK) subfamily in angiosperms. *Plant Physiology*, 170, 1595–1610. <https://doi.org/10.1104/pp.15.01470>
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, 40, D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
- Hanada, K., Zou, C., Lehti-Shiu, M. D., Shinozaki, K., & Shiu, S. H. (2008). Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiology*, 148, 993–1003. <https://doi.org/10.1104/pp.108.122457>
- He, J., Zhuang, Y., Li, C., Sun, X., Zhao, S., Ma, C., Lin, H., & Zhou, H. (2021). SIMP1 modulates salt tolerance by elevating ERAD efficiency through UMP1A-mediated proteasome maturation in plants. *New Phytologist*, 232, 625–641. <https://doi.org/10.1111/nph.17628>
- Hok, S., Danchin, E. G. J., Allasia, V., Panabières, F., Attard, A., & Keller, H. (2011). An *Arabidopsis* (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. *Plant, Cell & Environment*, 34, 1944–1957. <https://doi.org/10.1111/j.1365-3040.2011.02390.x>
- Hubert, P., Sawma, P., Duneau, J.-P., Khao, J., Hénin, J., Bagnard, D., & Sturgis, J. (2010). Single-spanning transmembrane domains in cell growth and cell-cell interactions. *Cell Adhesion & Migration*, 4, 313–324. <https://doi.org/10.4161/cam.4.2.12430>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14, 587–589.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kosuta, S., Held, M., Hossain, M. S., Morieri, G., MacGillivray, A., Johansen, C., Antolin-Llovera, M., Parniske, M., Oldroyd, G. E. D., Downie, A. J., Karas, B., & Szczygłowski, K. (2011). Lotus japonicus symRK-14 uncouples the cortical and epidermal symbiotic program. *The Plant Journal*, 67, 929–940. <https://doi.org/10.1111/j.1365-313X.2011.04645.x>
- Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. *Journal of Molecular Biology*, 305, 567–580. <https://doi.org/10.1006/jmbi.2000.4315>
- Le, M. H., Cao, Y., Zhang, X.-C., & Stacey, G. (2014). LIK1, a CERK1-interacting kinase, regulates plant immune responses in *Arabidopsis*. *PLoS ONE*, 9, e102245. <https://doi.org/10.1371/journal.pone.0102245>
- Lehti-Shiu, M. D., Zou, C., Hanada, K., & Shiu, S.-H. (2009). Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiology*, 150, 12–26. <https://doi.org/10.1104/pp.108.134353>
- Lemmon, Z. H., Reem, N. T., Dalrymple, J., Soyk, S., Swartwood, K. E., Rodriguez-Leal, D., Van Eck, J., & Lippman, Z. B. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nature Plants*, 4(10), 766–770.
- Li, E., & Hristova, K. (2010). Receptor tyrosine kinase transmembrane domains. *Cell Adhesion & Migration*, 4, 249–254. <https://doi.org/10.4161/cam.4.2.10725>
- Liu, P.-L., Du, L., Huang, Y., Gao, S.-M., & Yu, M. (2017). Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. *BMC Evolutionary Biology*, 17, 47. <https://doi.org/10.1186/s12862-017-0891-5>
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49, D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Mott, G. A. A., Thakur, S., Smakowska, E., Wang, P. W. P. W., Belkhadir, Y., Desveaux, D., & Guttman, D. S. D. S. (2016). Genomic screens identify a new phyto-bacterial microbe-associated molecular pattern and the cognate *Arabidopsis* receptor-like kinase that mediates its immune elicitation. *Genome Biology*, 17, 98. <https://doi.org/10.1186/s13059-016-0955-7>
- Munch, D., Gupta, V., Bachmann, A., Busch, W., Kelly, S., Mun, T., & Andersen, S. U. (2018). The Brassicaceae Family displays divergent, shoot-skewed NLR resistance gene expression. *Plant Physiology*, 176, 1598–1609. <https://doi.org/10.1104/pp.17.01606>
- Ngou, B. P. M., Heal, R., Wyler, M., Schmid, M. W., & Jones, J. D. G. (2022). Concerted expansion and contraction of immune receptor gene repertoires in plant genomes. *Nature Plants*, 8(10), 1146–1152.
- Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274. <https://doi.org/10.1093/molbev/msu300>
- Nimchuk, Z. L., Zhou, Y., Tarr, P. T., Peterson, B. A., & Meyerowitz, E. M. (2015). Plant stem cell maintenance by transcriptional cross-regulation of related receptor kinases. *Development*, 142, 1043–1049. <https://doi.org/10.1242/dev.119677>
- Pagnussat, G. C., Yu, H. J., Ngo, Q. A., Rajani, S., Mayalagu, S., Johnson, C. S., Capron, A., Xie, L. F., Ye, D., & Sundaresan, V. (2005). Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development*, 132, 603–614. <https://doi.org/10.1242/dev.01595>
- Panchy, N., Lehti-Shiu, M., & Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiology*, 171, 2294–2316. <https://doi.org/10.1104/pp.16.00523>
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, 5, e9490. <https://doi.org/10.1371/journal.pone.0009490>
- Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., Zhang, S., & Paterson, A. H. (2019). Gene duplication and evolution in



- recurring polyploidization-diploidization cycles in plants. *Genome Biology*, 20, 1–23.
- Renny-Byfield, S., & Wendel, J. F. (2014). Doubling down on genomes: Polyploidy and crop plants. *American Journal of Botany*, 101, 1711–1725. <https://doi.org/10.3732/ajb.1400119>
- Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Rizzon, C., Ponger, L., & Gaut, B. S. (2006). Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and rice. *PLoS Computational Biology*, 2, e115. <https://doi.org/10.1371/journal.pcbi.0020115>
- Robatzek, S., & Somssich, I. E. (2002). Targets of at WRKY6 regulation during plant senescence and pathogen defense. *Genes & Development*, 16, 1139–1149. <https://doi.org/10.1101/gad.222702>
- Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., & Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell*, 171, 470–480.e8. <https://doi.org/10.1016/j.cell.2017.08.030>
- Rodríguez-Leal, D., Xu, C., Kwon, C. T., Soyars, C., Demesa-Arevalo, E., Man, J., Liu, L., Lemmon, Z. H., Jones, D. S., Van Eck, J., Jackson, D. P., Bartlett, M. E., Nimchuk, Z. L., & Lippman, Z. B. (2019). Evolution of buffering in a genetic circuit controlling plant stem cell proliferation. *Nature Genetics*, 51(5), 786–792.
- Sakamoto, T., Deguchi, M., Brustolini, O. J., Santos, A. A., Silva, F. F., & Fontes, E. P. (2012). The tomato RLK superfamily: Phylogeny and functional predictions about the role of the LRR-RLK subfamily in antiviral defense. *BMC Plant Biology*, 12, 229. <https://doi.org/10.1186/1471-2229-12-229>
- Schallus, T., Fehér, K., Sternberg, U., Rybin, V., & Muhle-Goll, C. (2010). Analysis of the specific interactions between the lectin domain of malectin and diglucosides. *Glycobiology*, 20, 1010–1020. <https://doi.org/10.1093/glycob/cwq059>
- Schallus, T., Jaech, C., Fehér, K., Palma, A. S., Liu, Y., Simpson, J. C., Mackeen, M., Stier, G., Gibson, T. J., Feizi, T., Pieler, T., & Muhle-Goll, C. (2008). Malectin: A novel carbohydrate-binding protein of the endoplasmic reticulum and a candidate player in the early steps of protein N-glycosylation. *Molecular Biology of the Cell*, 19, 3404–3414. <https://doi.org/10.1091/mbc.e08-04-0354>
- Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J. P., Sun, L., Turner, S., & Karsch-Mizrachi, I. (2020). NCBI taxonomy: A comprehensive update on curation, resources and tools. *Database*, 2020. <https://doi.org/10.1093/database/baaa062>
- Shiu, S.-H., Karlowski, W. M., Pan, R., Tzeng, Y.-H., Mayer, K. F. X., & Li, W.-H. (2004). Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *The plant cell Online*, 16, 1220–1234. <https://doi.org/10.1105/tpc.020834>
- Shiu, S.-H. H., & Blecker, A. B. (2001). Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 10763–10768. <https://doi.org/10.1073/pnas.181141598>
- Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., Sankoff, D., DePamphilis, C. W., Wall, P. K., & Soltis, P. S. (2009). Polyploidy and angiosperm diversification. *American Journal of Botany*, 96, 336–348. <https://doi.org/10.3732/ajb.0800079>
- Tang, J., Han, Z., Sun, Y., Zhang, H., Gong, X., & Chai, J. (2015). Structural basis for recognition of an endogenous peptide by the plant receptor kinase PEPR1. *Cell Research*, 25, 110–120. <https://doi.org/10.1038/cr.2014.161>
- Tang, P., Zhang, Y., Sun, X., Tian, D., Yang, S., & Ding, J. (2010). Disease resistance signature of the leucine-rich repeat receptor-like kinase genes in four plant species. *Plant Science*, 179, 399–406. <https://doi.org/10.1016/j.plantsci.2010.06.017>
- Toufighi, K., Brady, S. M., Austin, R., Ly, E., & Provart, N. J. (2005). The Botany Array Resource: E-Northern, expression angling, and promoter analyses. *The Plant Journal*, 43, 153–163. <https://doi.org/10.1111/j.1365-313X.2005.02437.x>
- Uemura, T., Hachisu, M., Desaki, Y., Ito, A., Hoshino, R., Sano, Y., Nozawa, A., Mujiono, K., Galis, I., Yoshida, A., Nemoto, K., Miura, S., Nishiyama, M., Nishiyama, C., Horito, S., Sawasaki, T., & Arimura, G. I. (2020). Soy and *Arabidopsis* receptor-like kinases respond to polysaccharide signals from *Spodoptera* species and mediate herbivore resistance. *Commun Biol*, 3, 224. <https://doi.org/10.1038/s42003-020-0959-4>
- Van de Weyer, A.-L., Monteiro, F., Furzer, O. J., Nishimura, M. T., Cevik, V., Witek, K., Jones, J. D. G., Dangl, J. L., Weigel, D., & Bemm, F. (2019). A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. *Cell*, 178, 1260–1272.e14. <https://doi.org/10.1016/j.cell.2019.07.038>
- Wang, J., Hu, T., Wang, W., Hu, H., Wei, Q., & Bao, C. (2019). Investigation of evolutionary and expression relationships in the function of the leucine-rich repeat receptor-like protein kinase gene family (LRR-RLK) in the radish (*Raphanus sativus* L.). *Scientific Reports*, 9, 6937. <https://doi.org/10.1038/s41598-019-43516-9>
- Yang, H., Wang, D., Guo, L., Pan, H., Yvon, R., Garman, S., Wu, H. M., & Cheung, A. Y. (2021). Malectin/Malectin-like domain-containing proteins: A repertoire of cell surface molecules with broad functional potential. *The Cell Surface*, 7, 100056. <https://doi.org/10.1016/j.tcsu.2021.100056>
- Yeh, Y.-H., Chang, Y.-H., Huang, P.-Y., Huang, J.-B., & Zimmerli, L. (2015). Enhanced *Arabidopsis* pattern-triggered immunity by overexpression of cysteine-rich receptor-like kinases. *Frontiers in Plant Science*, 6, 322.
- Yeh, Y.-H., Panzeri, D., Kadota, Y., Huang, Y.-C., Huang, P.-Y., Tao, C.-N., Roux, M., Chien, H.-C., Chin, T.-C., Chu, P.-W., Zipfel, C., & Zimmerli, L. (2016). The *Arabidopsis* malectin-like/LRR-RLK IOS1 is critical for BAK1-dependent and BAK1-independent pattern-triggered immunity. *Plant Cell*, 28, 1701–1721. <https://doi.org/10.1105/tpc.16.00313>
- Yuan, N., Yuan, S., Li, Z., Zhou, M., Wu, P., Hu, Q., Mendu, V., Wang, L., & Luo, H. (2018). STRESS INDUCED FACTOR 2, a leucine-rich repeat kinase regulates basal plant pathogen defense. *Plant Physiology*, 176, 3062–3080. <https://doi.org/10.1104/pp.17.01266>
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J. D. G., Boller, T., & Felix, G. (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts agrobacterium-mediated transformation. *Cell*, 125, 749–760. <https://doi.org/10.1016/j.cell.2006.03.037>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kileeg, Z., Haldar, A., Khan, H., Qamar, A., & Mott, G. A. (2023). Differential expansion and retention patterns of LRR-RLK genes across plant evolution. *Plant Direct*, 7(12), e556. <https://doi.org/10.1002/pld3.556>