

*Annual Review of Plant Biology***Origin and Diversity of Plant
Receptor-Like Kinases**

Anne Dievart,^{1,2,*} Céline Gottin,^{1,2,*}
Christophe Périn,^{1,2} Vincent Ranwez,²
and Nathalie Chantret²

¹CIRAD, UMR AGAP, F-34398 Montpellier, France; email: anne.dievart@cirad.fr

²AGAP, Université de Montpellier, CIRAD, INRA, Montpellier SupAgro,
F-34060 Montpellier, France

**ANNUAL
REVIEWS CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Plant Biol. 2020. 71:131–56

First published as a Review in Advance on
March 18, 2020

The *Annual Review of Plant Biology* is online at
plant.annualreviews.org

<https://doi.org/10.1146/annurev-arplant-073019-025927>

Copyright © 2020 by Annual Reviews.
All rights reserved

*These authors contributed equally to this article

Keywords

receptor, kinase, plant, diversity, origin, domain, RLK, RK

Abstract

Because of their high level of diversity and complex evolutionary histories, most studies on plant receptor-like kinase subfamilies have focused on their kinase domains. With the large amount of genome sequence data available today, particularly on basal land plants and Charophyta, more attention should be paid to primary events that shaped the diversity of the RLK gene family. We thus focus on the motifs and domains found in association with kinase domains to illustrate their origin, organization, and evolutionary dynamics. We discuss when these different domain associations first occurred and how they evolved, based on a literature review complemented by some of our unpublished results.

Contents

1. INTRODUCTION	132
2. A SHORT HISTORY OF THE DISCOVERY AND DIVERSITY OF RECEPTOR-LIKE KINASE SUBFAMILIES	133
2.1. Receptor-Like Kinases with Diverse Extracellular Domains	133
2.2. Other Extracellular Motifs with Few Representatives or Uncommon Associations	139
2.3. Undefined Motifs in Extracellular Domain or No Associated Domain: RKF3 and LRK10-Type	140
2.4. Cytoplasmic or Membrane-Associated Receptor Kinases	140
3. THE ORIGIN OF THE RECEPTOR-LIKE KINASE DOMAIN	140
4. AT THE ORIGIN OF EXTRACELLULAR LIGAND-BINDING DOMAIN-KINASE DOMAIN ASSOCIATION	142
4.1. Receptor-Like Kinases in Angiosperms	145
4.2. Land Plant Clade	146
4.3. Charophyta Lineages	146
4.4. Chlorophyta Clade	147
5. CONCLUDING REMARKS	147

1. INTRODUCTION

Protein phosphorylation is a biochemical posttranslational modification involved in all signaling pathways and cellular activities in living organisms. It influences cell signaling networks in response to cellular or environmental stimulation, e.g., by the reversible regulation of protein functions through their activation or inactivation, the formation of protein complexes, and the arrangement of subcellular protein locations (31). The phosphorylation process, involving the transfer of a phosphoryl group from adenosine triphosphate (ATP) (or other nucleoside phosphates) to an acceptor hydroxyl residue of the protein substrate, is one of the roles ascribed to well-conserved protein kinase domains (48). Note that serine/threonine (Ser/Thr) phosphorylation is the most common phosphorylation event. Besides their role in this catalytic function, kinase domains also facilitate the binding and orientation of the ATP phosphate donor and protein substrate.

Networks of protein kinases involved in substrate phosphorylation are highly complex and consist of hundreds of proteins, making them one of the largest protein families (95). In 2002, Manning and coauthors (97) proposed the term kinome to encompass the complete set of protein kinases encoded in a genome. The kinome of the first plant genome sequenced in 2000 (*Arabidopsis thaliana*) included more than 1,000 proteins (2). Several years later, with more genomes available, the same analysis revealed that the plant kinome superfamily represented ~1–4% of the protein coding genes in 16 genomes (77). In plants, among the variety of protein kinases, the largest subfamily, besides those with functional homologs in other kingdoms (~400 proteins), consists of singular proteins with structural organization resembling that of receptor tyrosine kinases (RTKs) in animals. In 1993, Walker (153) introduced the term receptor-like kinases (RLKs) to account for some of the new genes he had cloned. This terminology is now used widely to define the whole subfamily of these plant receptors. Recent publications have also begun to name those that have a well-defined receptor function receptor kinases (RKs).

ATP: adenosine triphosphate

Ser/Thr: serine/threonine

RTK: receptor tyrosine kinase

RLK: receptor-like kinase

RK: receptor kinase

The first to promote the concept of receptor in the late nineteenth and early twentieth centuries were the immunologist Ehrlich and the physiologist Langley (94). Working on toxins and ferments, Ehrlich suggested the existence of molecules in body cells—initially referred to as side-chain and then receptor molecules—that need to fix substances in order to be biologically active (126). In parallel, another receptor theory introduced by Langley to explain antagonistic drug action in the body proposed the existence of receptive substances with chemical affinities for specific compounds (93). The questioning and controversy about the receptor theory ended in 1948 when the pharmacologist Ahlquist (1) published a paper describing two types of adrenergic receptors. The fundamental nature and structure of these receptors remained hypothetical until the 1980s. At this time, the first structurally related receptors from the RTK family were cloned. They are specific cell surface receptors for several growth factors, such as epidermal growth factor (EGF) or insulin, and their protein kinase activity was found to be intrinsic to receptor polypeptides and not part of a separate effector system (39, 145, 146). All of these proteins contained a large glycosylated extracellular ligand-binding domain (ECD), a hydrophobic domain that appeared to constitute a typical transmembrane (TM) domain, and an internal kinase catalytic domain (118). In plants, the first RLK, which is structurally similar to these RTKs, was discovered only 5 years later in 1990 (154).

EGF: epidermal growth factor
ECD: extracellular domain
TM: transmembrane
cDNA: complementary DNA
AA: amino acid
SP: signal peptide

2. A SHORT HISTORY OF THE DISCOVERY AND DIVERSITY OF RECEPTOR-LIKE KINASE SUBFAMILIES

It was assumed for decades that most cell-to-cell communication in plants would occur via cytoplasmic bridges called plasmodesmata because of the presence of cell walls separating plant cells from one another. Publication of the paper by Walker & Zhang (154) was thus a landmark event since it introduced a new paradigm for the mechanism by which cells communicate in plants. Although structurally resembling animal RTKs, plant RLK subfamilies underwent highly complex evolutionary histories due to considerable gene expansions (an increased number of genes per genome) coupled with diversifications (an increased number of specific combinations of domains) (57, 77, 79). The large number of domains found associated with kinase domains in these RLKs reflects the wide array of all cell communications that plants must handle within themselves and with the outside world, e.g., as pattern recognition receptors during immunity (26, 51, 139).

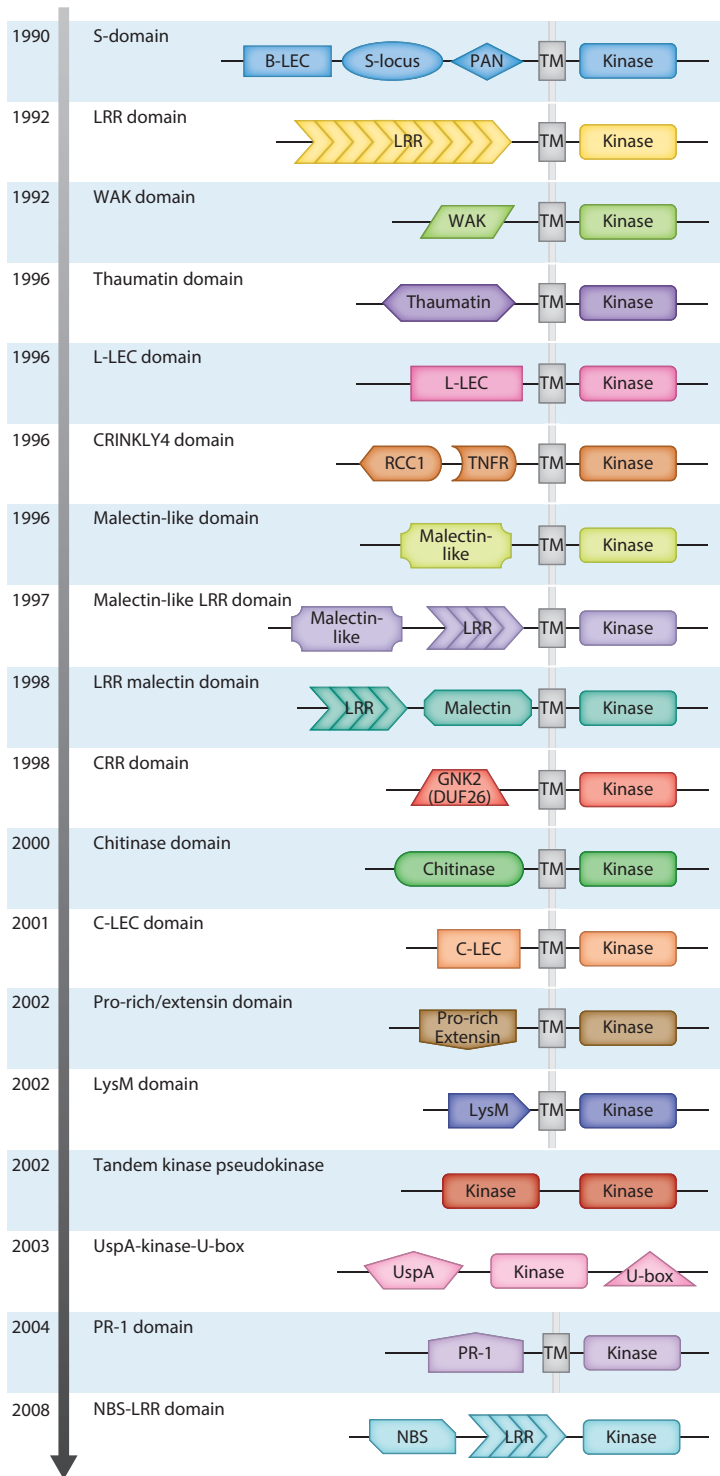
2.1. Receptor-Like Kinases with Diverse Extracellular Domains

In this section, we pay tribute to the pioneering researchers who discovered the first members of each of the new structural organizations observed in these receptors. Through them, the golden age for the discovery of hundreds of other RLKs in plants and the elucidation of their functions began (**Figure 1**).

2.1.1. S-domain (G-type lectin, B lectin). The first plant RLK was described in 1990, when Walker & Zhang (154) amplified and cloned, from a complementary DNA (cDNA) library of maize roots, the *ZmPK1* gene. This open reading frame, representing an almost complete transcript, encoded an 817–amino acid (AA) protein. This protein contained two hydrophobic regions, an N-terminal signal peptide (SP) of 28 AAs and a TM domain from residues 473 to 498, followed by a stop transfer sequence responsible for the interruption of the translocation process through the membranes. A comparison with the protein kinase known at this time in animals (48), and particularly the presence in *ZmPK1* of certain key residues that are highly conserved in the protein kinase catalytic domain, suggested that the *ZmPK1* receptor could be classified as a

Figure 1

Chronological overview of the major structural motif and domain organizations discovered in plant RLK proteins. Some of these domain organizations are found only in association with specific kinase subgroups. The relationships between these domain organizations and the group types classified according to the kinase phylogeny proposed by Lehti-Shiu & Shiu (77) are provided in **Supplemental Figure 1**. Abbreviations: B-LEC, bulb-type lectin; C-LEC, calcium-dependent lectin; CRR, cysteine-rich repeat; DUF26, DOMAIN OF UNKNOWN FUNCTION 26; GNK2, ginkbilobin2; LEC, lectin; L-LEC, legume lectin; LRR, leucine-rich repeat; LysM, lysin motif; NBS, nucleotide-binding site; PAN, plasminogen/hepatocyte growth factor family/apple domains of the plasma prekallikrein/coagulation factor XI family/ domains of various nematode proteins (142); PR-1, pathogenesis-related protein 1; RCC1, regulator of chromosome condensation 1; RLK, receptor-like kinase; TM, transmembrane; TNFR, tumor necrosis factor receptor; UspA, universal stress protein A; WAK, wall-associated kinase.



Ser/Thr kinase. This kinase domain would be intracellular and located downstream of the TM domain. Moreover, in the Ser/Thr kinase subfamily, the closest relative to ZmPK1 would be the human c-raf1 kinase (11), a proto-oncogene related to rapidly accelerated fibrosarcoma (RAF) Ser/Thr kinases, which were initially encountered in viruses (114, 165). The S-domain, in the extracellular part of the protein, has been subdivided into three subdomains: The first, the bulb-type lectin (B-LEC) domain, contains a highly conserved 14-residue stretch, and the third, the cysteine-rich (or PAN) domain, comprises 10–13 cysteine residues clustered near the TM domain (112, 142). Unexpectedly, besides these remarkable residues, the whole ECD showed 52% identity with the ECD of an S-locus-specific glycoprotein (SLG) of *Brassica* (104). These SLG genes were known to be among the products of the S-locus that controls the pollen-stigma interaction of self-incompatibility in the Brassicaceae mustard family. Moreover, even though the maize and *Brassica* genes shared some sequence similarities, data showed that they were expressed in different tissues, suggesting that ZmPK1 would not play a role in pollen-stigma recognition.

Screening cDNA libraries by using protein kinase genes as hybridization probes under low-stringency conditions was widely used in the 1990s to reveal serendipitously many new RLKs in several plant species. In 1991, Nasrallah's group (134) described the cloning of a new S-domain receptor kinase. The ECD displayed genotype-specific sequence polymorphisms that paralleled those of SLGs and S receptor kinase (SRK) transcripts were detected only in reproductive organs, suggesting that self-recognition between pollen and stigma during pollination in *Brassica oleracea* is mediated by receptor-ligand interactions between pollen and pistil components, leading to pollen acceptance or rejection. In 1992, the same team described the receptor ARK1 (141). Just like ZmPK1, ARK1 was expressed mainly in leaves, ruling out a role in self-incompatibility. Since then, several S-domain receptor kinases have been characterized, and besides their role in reproduction, some play a role in innate immunity, like mediating low-complexity bacterial metabolite sensing in *Arabidopsis* (71), in plant-mycorrhizal interaction (73), or in environmental and developmental processes (8, 109, 110, 117). These G-type lectin receptors [for *Galanthus nivalis* agglutinin (GNA)] belong to the large lectin (carbohydrate-binding) domain RLK family with other subfamilies, such as the L-type (L-LEC), C-type (C-LEC), malectin-like (also known as CrRLK1L), cysteine-rich repeat (CRR), and lysin motif (LysM) RLKs described below.

2.1.2. Leucine-rich repeat domain. In 1991, Bleeker (9) reported the cloning of an *A. thaliana* putative TM protein kinase. One year later, the *TRANSMEMBRANE KINASE 1* (TMK1) gene was described in detail by Chang and coauthors (21). In contrast with the previously cloned receptors, the ECD of TMK1 was composed of 11 copies of a 22-AA leucine-rich repeat (LRR) unit, a protein structural motif rich in leucine arranged in tandem repeats. In animals, the Toll receptor, involved in the establishment of dorsoventral polarity in *Drosophila* embryos, most resembled this new plant LRR receptor (50). The structural difference between these two receptors was the presence of the kinase domain in the cytoplasmic region of the plant receptor, which is absent in the Toll receptor [as it contains instead a Toll/interleukin-1 receptor (TIR) domain]. Since the 1990s, there has been enormous progress in determining the functions and mechanisms of activation of some of these receptors. Many functions have been attributed to these receptors, from plant growth and development to symbiosis and immunity. Activation of these receptors is generally based on peptide or hormonal ligand-dependent interactions of LRR receptor heterodimers containing one LRR-RLK with a few LRR units playing the role of coreceptor, such as BRASSINOSTEROID INSENSITIVE1-ASSOCIATED KINASE1 (BAK1) or SUPPRESSOR OF BIR1-1 (SOBIR) (46, 57, 58, 131, 160). However, because the LRR-containing receptor kinase subfamily contains the largest number of genes in all plant genomes studied so far (Table 1), many of these genes are still uncharacterized, despite efforts to uncover biologically

SLG: S-locus-specific glycoprotein

SRK: S receptor kinase

LRR: leucine-rich repeat

Table 1 Average number of proteins per species in monocots and dicots (when present), and percentage of species in which these domains are found

Domains	Average number of proteins per species in monocots (% of species)	Average number of proteins per species in dicots (% of species)
S-domain kinase	51 (96%)	54 (100%)
LRR kinase	189 (100%)	227 (100%)
WAK kinase	62 (100%)	42 (100%)
L-LEC kinase	44 (100%)	36 (100%)
CRINKLY4 kinase	1 (89%)	2 (97%)
Malectin-like kinase (CrRLK1L)	22 (96%)	30 (100%)
Malectin-like LRR kinase	9 (87%)	10 (100%)
LRR malectin kinase	3 (89%)	9 (97%)
CRR kinase	34 (96%)	37 (100%)
Pro-rich/extensin kinase	20 (96%)	14 (100%)
LysM kinase	10 (94%)	5 (100%)
C-LEC kinase	1 (91%)	1 (89%)
UspA-kinase-U-box	6 (96%)	9 (100%)
Thaumatococin	2 (64%)	1 (27%)
Chitinase	1 (60%)	2 (62%)
Kinase pseudokinase	6 (92%)	8 (96%)
PR-1 kinase	1 (30%)	1 (5%)

Abbreviations: C-LEC, calcium-dependent lectin; CRR, cysteine-rich repeat; CrRLK1L, *Catharanthus roseus* receptor-like kinase 1-like; L-LEC, L-type lectin; LRR, leucine-rich repeat; LysM, lysin motif; PR-1, pathogenesis-related protein 1; UspA, universal stress protein A; WAK, wall-associated kinase.

relevant interactions within LRR receptors containing a kinase domain or not, the latter being defined as receptor-like protein (RLP) (46, 82, 86, 128).

2.1.3. Wall-associated kinase domain. In 1992, Kohorn and coauthors (65) described a third type of RLK containing several EGF repeats in its ECD. This new protein, PRO25, was renamed WAK1 (for WALL-ASSOCIATED KINASE 1) a few years later (53). Recent discoveries have shown that these receptors are involved in cell wall integrity maintenance mechanisms through the perception of pectin or pectin-derived molecules not only during developmental cell wall extensions but also during stress- or pathogen-induced cell wall damage (5, 64).

2.1.4. Legume lectin domain. In 1996, Herve, Lesclure, and coauthors (55) described a protein, *Arabidopsis thaliana* lectin-receptor kinase 1 (Ath.lecRK1), containing an ECD homologous to the carbohydrate-binding proteins of the legume lectin family. Since 1889, extracts of castor beans or many other leguminous seeds were considered capable of agglutinating animal red blood cells via some remarkable proteins called lectins (from the Latin word *legere* for choose) (87, 135). However, despite this homology, all AA residues forming the putative monosaccharide-binding site of the lectin domain of Ath.lecRK1 were different from those usually found in legume lectins, suggesting that these receptors might not possess true lectin activity (14). To date, most studied L-LEC-RLKs have roles in plant immunity, with recent data suggesting that these receptors could bind metabolites such as extracellular adenosine triphosphate (eATP) (24) or extracellular nicotinamide

RLP: receptor-like protein

adenine dinucleotide (eNAD) (155). This latter molecule could be a key component of systemic acquired resistance in *Arabidopsis*, and the binding of eNAD to the L-LEC-RLK requires that this receptor forms a complex with the LRR-RLK BAK1 coreceptor (155, 157).

2.1.5. CRINKLY4 (and tumor necrosis factor receptor-like) domain. The 1995 Cold Spring Harbor Symposium review of Dangl, Preuss, and Schroeder (29) reports prime scientific discoveries including the cloning and functional analyses by Becraft, Stinard, and McCarty (7) of the maize receptor CRINKLY4. The ECD of this RLK was described as containing a novel 37-AA domain repeated seven times (the crinkly domain related to a GTPase-activating protein found in all eukaryotes) and a 26-AA region with similarities to the tumor necrosis factor receptor (TNFR) cysteine-rich region. The fact that this TNFR motif was known to contact the TNF suggests that the ligand for the maize protein might be a peptide. Even if a peptide of the CLE family has been suggested to be a ligand for the ARABIDOPSIS CRINKLY4 (ACR4) receptor in complex with the CLAVATA1 receptor belonging to the LRR-RLK subfamily, no biochemical data supporting this suggestion are available at the moment (28). It has to be noted that, contrary to previously described receptors that are present at several tens of copies per genomes, the CRINKLY4 subfamily is only represented by one or two copies on average per genome (**Table 1**).

2.1.6. Malectin-like (CrRLK1-like) domain. Schulze-Muth, Schröder, and coauthors (119) cloned the first member of the *Catharanthus roseus* receptor-like kinase 1-like (CrRLK1L) subfamily in 1996 during a screen of a library from a Madagascar periwinkle cell culture. None of the previously described ECDs showed similarity to the 450-AA residues composing the N-terminal part of this receptor. In 2011, Boisson-Dernier, Kessler, and Grossniklaus (10) showed that the ECD of these CrRLK1L receptors is homologous to a carbohydrate-binding domain, the malectin-like domain, which seems however to lack residues important for carbohydrate-rich ligand binding (37, 103). Various members of this RLK subfamily, in complex with other CrRLK1L receptors and glycosylphosphatidylinositol-anchored proteins chaperons/coreceptors, are involved in developmental or immune cell-wall-sensing responses induced by RAPID ALKALINIZATION FACTOR (RALF) peptides (43, 107, 161).

2.1.7. Malectin-like leucine-rich repeat domain. A malectin-like domain has also been found in the extracellular part of the receptors in association with LRRs. The light-repressible receptor protein kinase (LRRPK) is the first receptor of this type cloned in 1997 by Deeken & Kaldenhoff (33). The presence of a malectin-like domain followed by a few LRRs in the ECD of an RLK was described for the first time for the IMPAIRED OOMYCETE SUSCEPTIBILITY1 (IOS1) receptor in 2011 by Hok, Keller, and coauthors (59). It has recently been established that the CrRLK1L FERONIA and IOS1 function as scaffolds to regulate LRR-RLK immune receptor complexes (133, 164). Another well-studied receptor with this organization is the SYMBIOSIS RECEPTOR-LIKE KINASE (SYMRLK), which was cloned in 2002 by Stracke, Parniske, and coauthors (136). For this receptor, named DOES NOT MAKE INFECTIONS2 (DMI2) in *Medicago truncatula*, the protein level paced by protection from proteasome-mediated degradation could be a regulator of bacteria–plant and fungi–plant symbioses (108, 151).

2.1.8. Leucine-rich repeat malectin domain. In the LRR malectin domain, LRRs are located upstream in the extracellular part of the receptor, and the first gene of this type was described in 1998 by Takahashi, Chua, and coauthors (137). The extracellular part of receptor-like kinase

in flowers 1 (RKF1) possesses 13 LRRs. The presence of the malectin domain, located between LRRs and the TM domain in this receptor, was discovered later (38). The primary sequence difference between the malectin-like and malectin domains suggests that these two domains could have distinct ligand affinities. To our knowledge, scarce data are available on these LRR malectin RLKs in the literature, except that one of these receptors conferred enhanced resistance to the fungal pathogen *Magnaporthe oryzae* when overexpressed in rice (84).

2.1.9. Cysteine-rich repeat domain. Takahashi et al. (137) cloned a second gene, *RKF2*. In this 617-AA protein, the ECD has two copies of a C-X8-C-X2-C motif, distinct from the cysteine-rich region of SLGs and SRKs but highly similar to fungal lectins, suggesting either a common origin or convergent evolution of these domains (147). A few years later, the receptors possessing these CRR motifs were named CRKs, for CRR RLKs (22). These motifs are also named GINKBILOBIN2 (GNK2) [e.g., in the InterPro database (32, 102)] or DOMAIN OF UNKNOWN FUNCTION 26 (DUF26) (in the literature). The latest research on this receptor subfamily suggests that they play roles in defense and oxidative stress responses, programmed cell death, and development (13, 75, 111, 147).

2.1.10. Pro-rich/extensin domain. In 2002, Silva & Goring (125) described and characterized a new gene from *Brassica napus*: *PROLINE EXTENSIN-LIKE RECEPTOR KINASE1 (PERK1)*. This gene possesses a proline-rich ECD with sequence similarity to the extensin family of proteins, and *PERK1* gene expression was rapidly induced by wounding, suggesting a role in sensing cell wall modifications following cell wall damage or pathogen responses, similar to the WAKs. Even though characterized *PERK* genes suggest that they have functions in development and virus infection, up to now, ligands for these receptors have not been described (12).

2.1.11. Lysin motif domain. In the search for receptors involved in the perception of Nod factors during the early steps of nodulation following rhizobial infection of legumes, a hypothesis on the presence of two receptors was put forward in 1994 (3). However, it took almost 10 years before two different teams, led by Stougaard (92, 113) and Geurts (88), published their results on the cloning of a yet-undescribed type of receptor containing LysMs in their ECD. The receptors analyzed in these papers, i.e., NOD FACTOR RECEPTOR1 (NFR1), NFR5, and SYM2 [now known as LysM domain-containing receptor-like kinases (LYKs)], contained two or three LysM modules that occur in bacterial peptidoglycan-binding proteins or in chitinases from yeast and alga (20). Based on studies conducted over the last decade, it is now clear that all LysM-RLK receptors characterized to date, in complex with LysM-RLP or with a kinase domain devoid of kinase activity [called non-RD kinases (30) and representing 20% of all kinases (72)], bind to chitooligosaccharides ligands (89). Their dual role in plant immunity and in the establishment of the arbuscular mycorrhizal and rhizobium-legume symbioses is also well-defined (19).

2.1.12. Calcium-dependent lectin domain. The association between the calcium-dependent (or C-type) lectin (C-LEC) motif and a kinase domain in plants was noted for the first time in *Arabidopsis* by Shiu & Bleeker (121) in 2001. Interestingly, this association was observed in a single receptor. Contrary to the abundance of other types of lectin receptor kinase in plants, the very low number of receptors with this C-LEC kinase structural organization has now been confirmed in several genomes (one or two copies) (8, 124, 148, 163). However, we still do not know their functions. In animals, C-LEC receptors are numerous and have been extensively studied because of their decisive role in pathogen recognition and immunity (16).

2.2. Other Extracellular Motifs with Few Representatives or Uncommon Associations

Several new domain associations have been discovered in the search for Resistance genes (R genes). The RLKs described in this section are representative of domain associations that are sometimes species-specific and present in only a few copies in plant genomes. Consequently, for many of them, more extensive investigations are needed to uncover their real functions and activation mechanisms.

PR:
pathogenesis-related

2.2.1. Chitinase (glycoside hydrolase)-type domain. In 2000, Kim, Pai, and coauthors (62) cloned the *CHITINASE-RELATED RLK1 (CHRK1)* gene. This new type of receptor contains a domain closely related to chitinase in its extracellular part, but lacking the essential residue required for chitinase activity. Co-suppression of the endogenous tobacco gene showed pleiotropic developmental phenotypes, and mRNA accumulation was strongly stimulated by fungi and viruses (76).

2.2.2. Thaumatin domain. In 1996, Wang, Lawton, and coauthors (156) published a paper stating that “The PR5K Receptor Protein Kinase from *Arabidopsis thaliana* Is Structurally Related to a Family of Plant Defense Proteins.” Indeed, the extracellular part of this pathogenesis-related (PR)-5-like RLK contains a thaumatin-like motif, which shares sequence similarities with thaumatin, a sweet-tasting protein originally found in an African berry (150). Thaumatin RLKs are transcriptionally induced by pathogenic and environmental stress, and some of these receptors play a specific role in abscisic acid-dependent drought-stress signaling (6, 144).

2.2.3. Tandem kinase-pseudokinase. In 2002, Brueggeman and coauthors (18) cloned the stem rust-resistance gene *Rpg1* in barley. This gene had homology to receptor kinases, but the domain organization was unique: Two kinase domains were fused into one protein, with the second one being a pseudokinase. Because this structure has been found in a wide range of plant taxa, it has recently been termed tandem kinase-pseudokinase (TKP) (63).

2.2.4. Nucleotide-binding site leucine-rich repeat kinase. In 2008, Brueggeman and coauthors (17) described the *Rpg5* gene, which is also involved in stem rust resistance. This time, this new R gene was found to code for a protein with a nucleotide-binding site (NBS), LRRs, and protein kinase domains. This organization is now defined as integrated domain (ID) nucleotide-binding leucine-rich repeat (NLR) proteins (NLR-IDs). NLR-IDs have been found in many plant species, and up to 10% of NLRs contain IDs (68, 115). The kinase association, as for other IDs, could be a strategy to trap kinase-targeting pathogen effectors (66; see also 138).

2.2.5. Pathogenesis-related protein 1 domain. In 2004, a new structural organization was described by Shiu and colleagues (123) in a paper comparing *Arabidopsis* and rice RLKs. The domain pathogenesis-related protein 1 (PR-1) was found to be fused to a kinase domain related to the CRR subfamily. Several years later, in 2013, the same ECD configuration was detected in two genes in the *Theobroma cacao* genome (140). The PR-1 domain was found to be fused to the TM and kinase domains related to the L-LEC subfamily, suggesting that these domain associations had occurred at least twice. The origin of this fusion in cacao seemed to involve retrotransposition since one of the genes was surrounded by retrotransposition marks and repetitive elements.

2.3. Undefined Motifs in Extracellular Domain or No Associated Domain: RKF3 and LRK10-Type

Takahashi et al. (137) described a third receptor, RKF3. It possessed all the features of a canonical receptor kinase, namely an SP, an ECD, a TM domain, and a kinase domain. However, in the 200 AAs forming the ECD, no distinct motifs were found. This was also the case for the LRK10 receptor cloned in wheat by Feuillet and coauthors (41). This receptor is part of a gene family, wheat leaf rust kinase (WLRK), containing three conserved regions separated by three variable regions in their ECD.

2.4. Cytoplasmic or Membrane-Associated Receptor Kinases

Besides these receptor structural configurations, some kinase proteins lacking an apparent TM domain but closely related to these RLKs were also discovered.

2.4.1. Kinase only. For example, the *Arabidopsis* APK1 protein, which was described as phosphorylating tyrosine, serine, and threonine, was cloned by Hirayama & Oka (56) in 1992. This protein is composed of 410 AAs and possesses only a kinase domain with a short N-terminal domain containing an N-myristoylation site, but no TM domain or ECD was identified. Moreover, in 1993, Martin, Tanksley, and coauthors (99) reported cloning the first kinase domain-containing protein conferring bacterial speck resistance in tomato. The cloned *Pto* gene encoded a 321-AA hydrophilic protein that lacked the TM domains and ECDs. As for APK1, the presence of a putative myristoylation site suggested that the *Pto* kinase could be membrane-associated but not an integral membrane protein. In 1999, Snyder & Kohorn (129) revealed other interesting *Arabidopsis* receptors associated with energy transduction. These thylakoid-associated kinase RLKs (TAK RLKs), with no ECD but with a single N-terminal hydrophobic region, were the first RLKs found to be involved in the chloroplast phosphorylation network. Up to now, the kinase-only proteins that have been studied function in complex with S-domain, LRR, malectin-like, and LysM RLKs. They are phospho-dependent transducers acting downstream of RLKs, and they intervene in all signaling pathways and functions fulfilled by RLKs (85).

2.4.2. Universal stress protein A-kinase-U-box domains. Finally, in 2003, Kerk, Gribskov, and coauthors (61) described a new structural configuration consisting of an N-terminal universal stress protein A (UspA) domain of *Escherichia coli*, a kinase domain related to RLKs, and a C-terminal U-box domain. The association of a kinase domain followed by a U-box domain had already been reported in 2001 in an article discussing the characteristics of a new protein family called plant U-box (PUB) proteins (4). To our knowledge, none of these proteins have been studied so far, but, with the ID model in mind, these proteins could also play the role of a trap in the many RLK signaling pathways in which PUBs and ubiquitination have been involved (27, 40). Alternatively, this could simply be a way to fuse two proteins that normally work subsequently.

This Prévert-style inventory of researchers and articles describing the discoveries of new RLKs in the last decade of the twentieth century has laid the foundations for showcasing that (a) these receptors are involved in a myriad of different functions such as developmental processes, disease resistance, symbiosis, or self-incompatibility, (b) they are present in many plant species, and (c) they are all phylogenetically related.

3. THE ORIGIN OF THE RECEPTOR-LIKE KINASE DOMAIN

The high diversity of the domain combinations in the RLKs we have described, especially in their ECD, raises several questions regarding their origin: How are RLK domains related to kinase

domains in other kingdoms? When did each of these combinations appear? Did they appear around the same time or at very distinct periods? Did each of them appear only once, or did some of them appear multiple times over the course of evolution? A journey into the past is warranted to answer these questions.

Hanks, Quinn, and Hunter (48) were the first to initiate, with the data set of protein kinases available in 1988, a classification based on similarities in their AA sequences. They thus aligned the 250–300 AAs of 65 protein kinase catalytic domains from vertebrates and invertebrates. This allowed them to precisely describe, for the first time, the 11 major conserved subdomains of the catalytic domains. Additionally, their phylogenetic analysis revealed five major clades that pooled protein kinases with similar modes of regulation or substrate specificities. One of the clades grouped protein tyrosine kinase sequences closely related to the Ser/Thr kinase Raf protein. In 1995, Hanks & Hunter (47) proposed a new classification of a larger set of 400 protein kinases, this time representative of all kingdoms. Among the 99 protein kinase catalytic domains used to infer the kinase phylogeny based on the maximum parsimony principle, 41 plant protein kinases were included in this tree, 6 of which were RLKs, 3 containing an S-domain [ZmPK1 (154), SRK2 and SRK6 (134)], 1 an LRR domain [TMK1 (21)], 1 a WAK domain [Pro25 (WAK1) (65)], and 1 with no associated domain [APK1 (56)]. The resulting tree presented in that paper revealed that the plant RLK sequences clustered together, suggesting a monophyletic origin. They were classified as “other protein kinase families (not falling in major groups).” Moreover, the closest animal kinase to this group was the *Drosophila* Pelle kinase, indicative of a possible common origin of plant RLKs and animal Pelle kinases. The fruit fly kinase was also classified in the “other protein kinase families (not falling in major groups)” category, but in a different subgroup. Another important result from this analysis is the fact that, in the S-domain group, ZmPK1 and SRKs were not grouped together in the inferred tree, even though they both contain an S-domain associated with the kinase domain.

The second attempt to build a phylogenetic tree combining RLK and animal kinase domains was published in 1997 by Clark, Williams, and Meyerowitz (25). In their phylogenetic tree, based on the minimum evolution principle, all plant RLKs analyzed [S-domain: ZmPK1, RLK1 and RLK4 (153), and SRK; LRR domain: CLAVATA1 (CLV1) (25), RLK5 (153), TMK1, ERECTA (ER) (143), and Xa21 (132); CRINKLY4 domain: CR4 (7); and no associated domain: PTO (99) and FEN (91)] formed separate lineages distinct from animal kinases. Moreover, the four S-domain receptors were again not clustered together, with SRK being outside of the RLK1, RLK4, and ZmPK1 clade, thus confirming previous results (47). Another observation was that the no associated domain receptors PTO and FEN clustered together and were close relatives to an LRR domain receptor clade, including CLV1, RLK5, and ER. However, some other LRR domain sequences, i.e., Xa21 and TMK1, were not included in this group, and formed two disjointed clades, also distinct from the CRINKLY4 domain CR4 clade. In 1999, Hardie (49) published a review in which the phylogenetic relationships of 89 kinase domains from *A. thaliana* were considered, including 18 RLKs. This analysis showed that these protein kinases clustered into 12 major subfamilies, one being RLKs. However, as noted in the article, the kinase domain sequences of RLKs were more variable than those of the other subfamilies. Moreover, these sequences did not cluster in exactly the same way, depending on whether the analyses were done using their kinase or ECD sequences. Four subclasses emerged in the phylogenetic tree inferred using the ECDs of RLKs. The first subclass is the S-domain containing RLK1, RLK4, and ARK1 (141). The second is the LRR domain containing CLV1, ER, RLK5, TMK1, TMKL1 (149), RKF1 (137), and the BRASSINOSTEROID-INSENSITIVE1 (BRI1) receptor (83). The third is the WAK domain containing Pro25 (WAK1) and WAK4 (52). The fourth is the L-LEC domain containing LecRK1 (55). The accumulation of *Arabidopsis* sequences available in databases by

Mya: million years ago

2000 (101), and the recent release of the complete sequence of the *Arabidopsis* genome, revealed that the *Arabidopsis* genome contains nearly 1,000 genes encoding Ser/Thr protein kinases, several hundred of which are RLKs (2). Comparative genome analysis between *Arabidopsis* and other sequenced genomes again highlighted that RLKs were highly similar to the *Drosophila* Pelle protein kinase and the mammalian INTERLEUKIN1 RECEPTOR-ASSOCIATED KINASE (IRAK).

Overall, these results provided a first indication that (a) RLKs have a monophyletic origin and are close relatives to animal Pelle kinases and IRAKs, (b) RLK kinase domains evolve more rapidly than kinase domains of other protein kinases, (c) some ECDs found in plants are unfamiliar in animals, and (d) although the ECDs are similar, kinase domains of some of these RLK receptors do not cluster together in the phylogenetic trees, suggesting that several associations between these ECDs (e.g., S-domain or LRR) and different kinase domains could have occurred independently. These results paved the way for Shiu & Bleecker (121) to publish a first remarkable paper in 2001. All of this evidence supports the monophyletic origin of the RLK/Pelle subfamily [renamed by Shiu & Bleeker (121)], and the large-scale expansion of this gene family in plants (few genes in animals versus hundreds in plants) (123). However, the time of divergence of this clade from other kinases has yet to be precisely established. It seems clear from these data that the origin of the RLK/Pelle clade could predate the divergence of plants and animals (121), i.e., around $1,580 \pm 90$ million years ago (Mya) (54), but the presence of RLK/Pelle in the Alveolata phylum [e.g., *Plasmodium* (122), *Perkinsus*, and *Toxoplasma* (79)] suggests that it may be even older, i.e., between 1,580 and 1,840 Mya (54) (**Figure 2**). Analyses of new genome sequences revealed, however, that RLK/Pelle orthologs are not found in the excavates *Leishmania major* and *Giardia intestinalis* (96), in *Dictyostelium*, the model organism of Amoebozoa (44), in fungal genomes, or in *Monosiga brevicollis*, a choanoflagellate close relative to animals (79, 121, 122), therefore suggesting massive losses in these lineages under this hypothesis. A search for RLK/Pelle sequences in many other species in these phyla should thus help to complete the evolutionary history of these kinase domains, which could be more complex than anticipated [e.g., involving horizontal transfers (79) or convergent evolution].

4. AT THE ORIGIN OF EXTRACELLULAR LIGAND-BINDING DOMAIN-KINASE DOMAIN ASSOCIATION

Two approaches are found in the RLK-related literature that are focused on either a specific RLK subfamily or a specific taxon. The first approach analyzes a given RLK subfamily, and it is studied among several species in order to reconstruct its evolutionary history. This approach and the resulting conclusions depend on the genomes available at the time of the analyses. The second approach focuses on a given species, and it aims to inventory all of its RLKs, generating a very precise picture of a given step in the evolution of the RLK family. Below we summarize findings on RLK evolution derived from the literature as well as our studies using these two approaches.

The seminal extensive genome-wide analyses conducted by Shiu & Bleecker (121, 122) provide an evolutionary framework that is highly useful for the whole scientific community working on RLKs. Moreover, additional later studies by Shiu and collaborators (77–79, 123) have also greatly helped gain further insight into the evolutionary dynamics of this gene family in plants. Indeed, these were the first reported large-scale analyses of RLK gene families comparing several species. Phylogenetic analyses based on kinase domains combined with annotations of ECD structural domains led to the first attempt to classify this large family of plant receptors. The RLK family was in turn subdivided into more than 50 different subfamilies, around 20 of which are LRR receptor kinase subfamilies and 20 are receptor-like cytoplasmic kinases (RLCKs) that have been defined as receptor kinases with no apparent SP or TM domain (grouping the kinase-only and the UspA kinase U-box) (121). These different subfamilies tend to have similar structural organization in

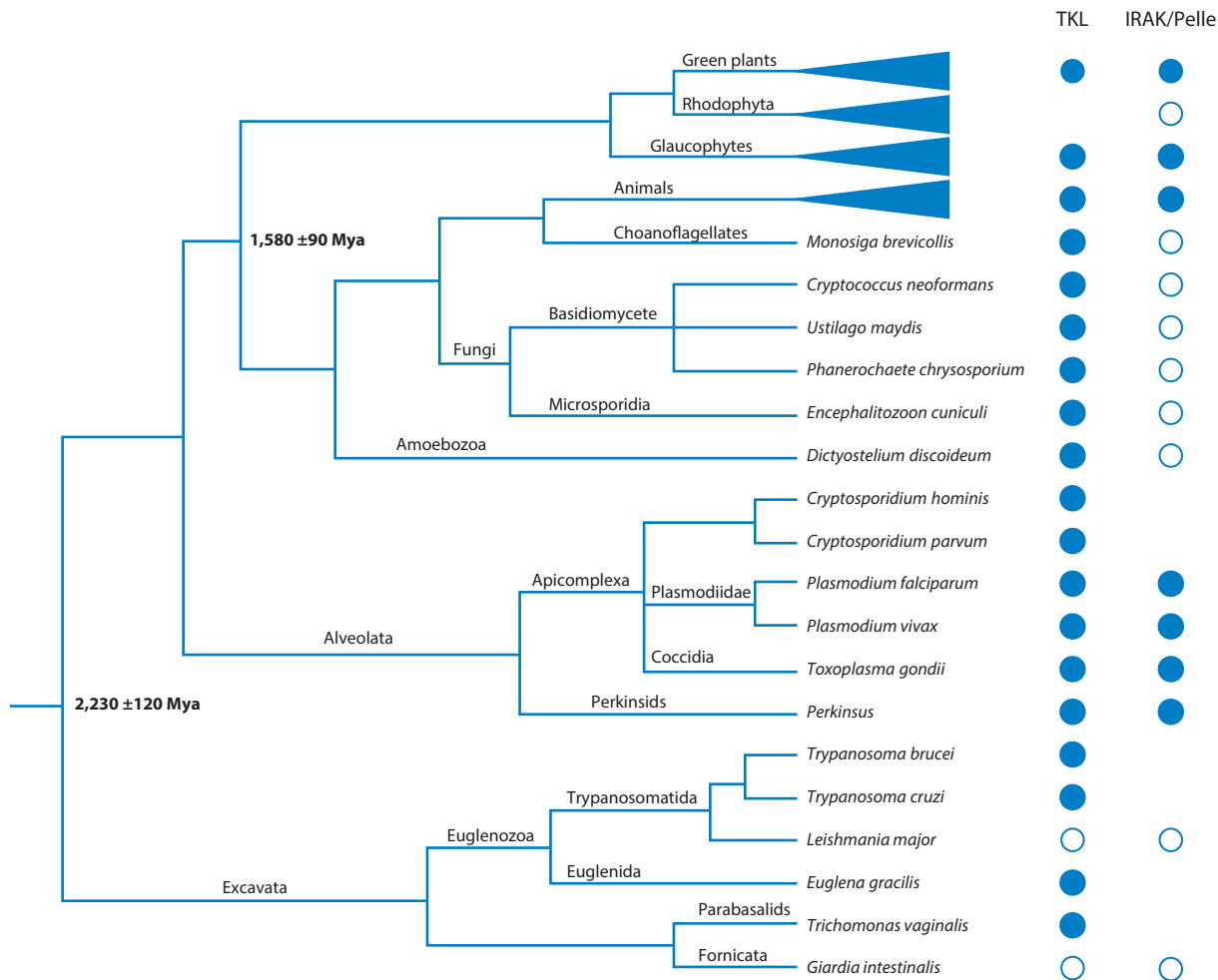


Figure 2

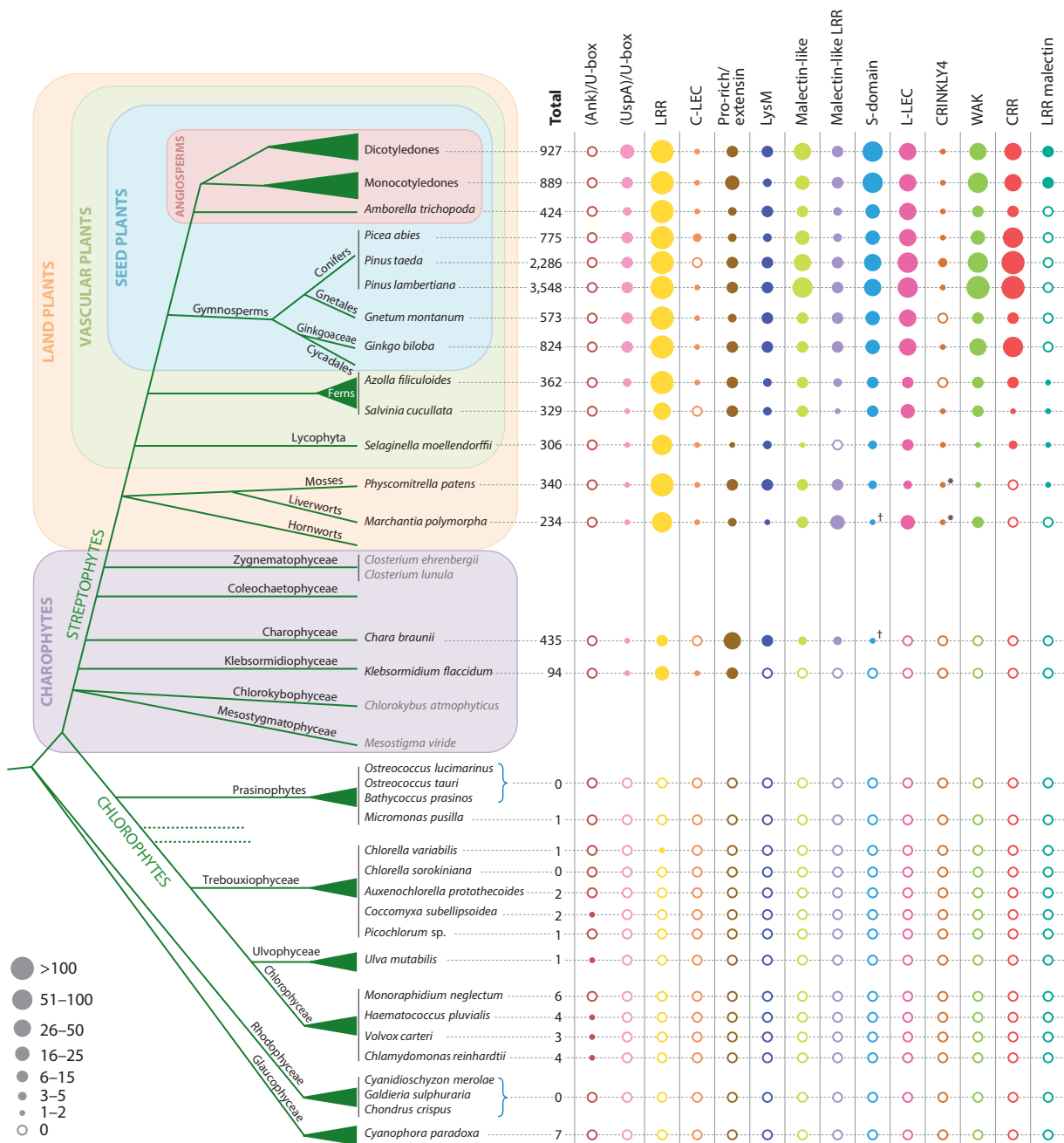
The presence/absence of tyrosine kinase-like (TKL) proteins and inter leukin 1 receptor-associated kinase (IRAK)/Pelle kinase families in the tree of life, showing a schematic representation of phyla and species in which TKL and/or IRAK/Pelle kinases have been found (filled circle), searched but not found (open circle), or not searched yet (blank). The tree is based on the Tree of Life Web Project (<http://tolweb.org/tree/>) and Reference 54. Note that based on Hanks classification, RLKs are part of the tyrosine kinase-like (TKL) IRAK/Pelle subgroup (45, 69, 98).

their ECDs (**Supplemental Figure 1**). This large body of literature has shed considerable light on the extraordinary gene expansion, i.e., the increasing number of genes in some subfamilies, as the consequence of duplication events, including whole genome, segmental, or tandem duplications.

In the streptophytes, the number of proteins within RLK subfamilies expanded, but also the number of these different subfamilies increased (resulting in diversification) (70). An assessment of the presence and structural architecture of RLK ECDs in the different lineages is required to trace the origins of RLK diversification, i.e., the increasing number of specific domain combinations. In the early twenty-first century, whole plant genome sequences have started to be released, and a wealth of data is now available (100). For example, the kinomes of almost 200 genomes of plants are stored in the iTAK database (166), and they are classified based on the latest Shiu

Supplemental Material >

classification proposed in 2012 (77). Moreover, the recent availability of many new genome sequences in Charophyta and basal land plant lineages has led to a more precise description of the fusion events that have shaped RLK subfamilies in angiosperms. These data from the literature and our unpublished studies are described below and summarized in **Figure 3**.



(Caption appears on following page)

Figure 3 (Figure appears on preceding page)

The presence and frequencies of RLK domains within major green lineages. The color code is the same as in **Figure 1**; a filled circle depicts the presence of the concerned structure, and its size is proportional to the number of sequences found; an open circle means that no gene has been found. For a brief description of the current knowledge on green plant lineage (e.g., Viridiplantae) phylogeny, readers are referred to **Supplemental Data 1**. There are two areas of discrepancy between published results and ours. The first, shown by an asterisk (*), is the appearance of the CRINKLY4 subfamily in *Marchantia* (15). Our data suggest that the fusion between the CRINKLY4 domain and a kinase appeared in the *Physcomitrella patens* genome (74) and not in *Marchantia*. Previous phylogenetic analysis of the CRINKLY4 subfamily showed that only one protein was a true CRINKLY4 family member containing all of the characteristic domains in the *P. patens* genome (105). This discrepancy could be due to the different levels of importance given to the presence of structural domains in the CRINKLY4 RLK. These results should then be taken with caution. The second discrepancy is shown by a dagger (†). Our data suggest that S-domain RLKs could be a new putative receptor configuration that appeared before the emergence of the Phragmoplastophyta clade (that includes three charophyte lineages: Charophyceae, Coleochaetophyceae, and Zygnematophyceae, together with land plants). However, these observations refer to a sole sequence that, in our study, would have been lost in *Marchantia*. Based on published data, the S-domain RLK subfamily only appeared in land plants (15, 162). Gray text indicates a species that is included in this article, but for which a complete proteome is not available. Abbreviations: ANK, ankyrin repeat; C-LEC, calcium-dependent lectin; CRR, cysteine-rich repeat; L-LEC, L-type lectin; LRR, leucine-rich repeat; LysM, lysin motif; RLK, receptor-like kinase; UspA, universal protein A; WAK, wall-associated kinase.

Supplemental Material >

4.1. Receptor-Like Kinases in Angiosperms

To supplement the data harvested in the literature, we selected 176,020 protein kinases classified as RLK/Pelle in the iTAK database (166) and scanned them to annotate their structural domains (**Supplemental Table 1**). On average, there are approximately 900 RLKs per genome in the 127 dicot and 53 monocot species analyzed in accordance with previous results (123, 167). One third consist only of a kinase domain (62,642 proteins). Some of these proteins are RLCKs. In 12% of the analyzed receptor configurations, only a TM and kinase domain were detected without any other motifs in the ECD (13,591 proteins). However, in many of these proteins, the ECD was very long (several hundred AAs) (**Supplemental Figure 2**). New motifs or domains may ultimately be discovered in these receptors.

More than 2,000 unique domain organizations were found in the protein data set. Among them, 1,730 were found in less than 10 proteins in all of the 180 genomes (**Supplemental Table 1**). Moreover, many variations per domain have been noted in multidomain ECDs (162) (see also **Supplemental Figure 3**). These variations shed light on the multiple domain organizations found in plant species, and show the trial by error dynamics within genome evolution. These diverse organizations also reveal that these domains have often been reused throughout evolution, e.g., the association of RLPs with kinase domains. Of course, many of these configurations are probably annotation errors (e.g., gene boundary delimitation leading to protein fusion prediction errors or a lack of domain recognition by the software) since these architectures are based on automatic genome annotations. Nevertheless, the 12 types of domain described in Section 2.1 are the most abundant predicted structures observed in RLKs in this data set (**Table 1**). The largest subfamilies are LRR and putative carbohydrate-binding [containing the S-domain, L-LEC, C-LEC, malectin, malectin-like (CrRLK1L), LysM (8) and CRR (147) kinases], with over 200 LRR and 150 putative carbohydrate-binding receptors per genome on average. Note that only one C-LEC RLK was present per genome, as mentioned above. Could the uniqueness of the C-LEC receptor in plant genomes be an indication of its functional importance? Among these ECD kinase associations, only one was probably not present in the angiosperm ancestors, i.e., the LRR malectin receptor kinase that is not found in the *Amborella* genome and whose expansion seems to have been monocot- and dicot-specific (38, 42) (**Figure 3**). Note, however, that the evolutionary history of this subfamily is perhaps more complex since some LRR malectin receptors—despite also being absent in gymnosperms—are found in basal vascular plants and in the *Physcomitrella* genome (one or two copies).

For the structures described in Section 2.2 (**Table 1**) and for many others (**Supplemental Table 1**), a sporadic distribution in monocot and dicot species has been observed without any obvious lineage-specific expansion. This rate of association is assumed to have been very low (78). If the events we noticed were not due to annotation errors, these domain swapping events could represent new randomly occurring domain associations. The putative ones that caught our attention are (a) the integration of a WAK domain in the S-domain organization, (b) the gain of a kinase domain in NBS-LRR proteins—but unlike Rpg5 (17), the kinase domain here is located at the N-terminal end with respect to the NB-ARC and LRR domains—and (c) receptor kinases containing only a malectin domain but no detected LRR repeats (**Supplemental Figure 4**). However, as these proteins have only been characterized by sequence analysis, further functional studies of these genes should provide more insight into the functions of these putative receptors.

4.2. Land Plant Clade

Prior to the publication of complete genomes of species belonging to the land plant group but outside of angiosperms, 29 cDNAs of RLKs (belonging to the LRR, C-LEC, Pro-rich/extensin, LysM, malectin-like (CrRLK1L), malectin-like LRR, and L-LEC subfamilies) were identified in *Marchantia polymorpha* (116). In the recent publication of the complete *M. polymorpha* genome (15), a whole section is devoted to the study of RLK genes, revealing the presence of many of them in this early diverging lineage of land plants. Although some subclasses were not represented (e.g., the BRI1 clade in the LRR subfamily), all domain RLK subfamilies were present except the CRR RLKs that appeared in vascular plants (*Selaginella*) (147) (see the caption for **Figure 3** for the discrepancy observed for the CRINKLY4 and S-domain subfamilies).

4.3. Charophyta Lineages

Before the complete genomes of charophyte algae were sequenced, a few studies explored the presence of RLKs in cDNA libraries, and some were found in two Charophycean algae, *Closterium ebrenbergii* (Zygnematophyceae, 14 RLKs) and *Nitella axillaris* (Charophyceae, 13 RLKs) (116). In another study, the presence of symbiotic gene homologs within the entire green lineage revealed that homologous sequences of LysM, malectin-like (CrRLK1L) and malectin-like LRR-domain RLKs were present in charophyte lineages, so these associations appeared before the emergence of land plants (34, 152) (**Figure 3**). The availability of the complete genome of *Chara braunii* revealed that the LysM RLK gene family was not only present but also expanded in this species (seven copies) (106). Since L-LEC, CRINKLY4, WAK, and putatively S-domain subfamilies are found in basal land plants but not in the *C. braunii* genome, the absence of these receptor configurations in Zygnematophyceae (23) and Coleochaetophyceae genomes would be highly indicative of their land plant specificity.

The earliest Streptophyta sequenced genome available today is *Klebsormidium flaccidum* (60). Although a detailed list of RLKs was not specifically established, RLKs were clearly identified as a gene family in *Klebsormidium*, for which the numbers of genes were significantly increased in land plant genomes (60). We found 94 RLKs in this genome, and the first canonical receptor configuration with an acquisition of a TM domain was observed in it. The only kinase-associated ECDs found in this genome are the LRR, C-LEC, and Pro-rich/extensin domains, thus suggesting that these fusions were the first that occurred and arose before the Klebsormidiophyceae split.

Mesostigmatophyceae (represented by *Mesostigma viride*) and Chlorokybophyceae (represented by *Chlorokybus atmophyticus*) are the earliest known diverging streptophyte lineages (81). So far, no RLKs have been identified from EST/cDNA data (127), but complete genomes in

these very early diverging lineages would be essential to delve deeper into the origin of the associations between LRR, C-LEC, and Pro-rich/extensin domains, which may have emerged early in the streptophyte lineage.

4.4. Chlorophyta Clade

In the Chlorophyta clade, there have been several attempts to screen for the presence of RLKs. A search in the *Ostreococcus tauri* (35) and *Chlamydomonas reinhardtii* genomes revealed the presence of two RLCK genes in *C. reinhardtii* but none in *O. tauri* (78). When using the standalone iTAK program (166) to assess several chlorophyte genomes, we found a few kinase domains classified as RLK/Pelle, some with various associated domains [e.g., ankyrin- or WD40-repeats annotated by the InterProScan program (102) but no TM domain detected by the TMHMM program (67)]. One domain that was found repeatedly in several species from the Chlorophyceae (three copies), Ulvophyceae (one copy), or Trebouxiophyceae (one copy) was a C-terminal U-box domain associated or not with other N-terminal domains resembling the UspA kinase U-box RLCKs described above (Figure 3). The presence of this particular assembly in some RLCK genes, also detected throughout the streptophyte lineage, could be a strong clue that this subfamily may have been present before the streptophyte-chlorophyte split, but after the Rhodophyceae divergence. Moreover, they could represent the ancestors from which RLKs arose.

Some LRR kinase domain associations have been observed in *Chlorella variabilis* (Trebouxiophyceae, Chlorellales) but not in any other chlorophytes analyzed previously [*C. reinhardtii* and *Volvox carteri* (Chlorophyceae), *Micromonas pusilla*, *Ostreococcus lucimarinus*, and *O. tauri* (prasinophytes)] (36, 90, 130). In the 14 species analyzed in this study, an LRR domain kinase association was again only detected in *C. variabilis*, and it was not found even in its close relative *Chlorella sorokiniana*. Previous phylogenetic analysis showed that this kinase domain did not cluster with the monophyletic RLK/Pelle subfamily (36). Overall, these results suggest that the LRR kinase association observed in this sole *Chlorella* genome could be a unique event that occurred independently of the successful association that spread in the streptophyte genomes.

5. CONCLUDING REMARKS

In conclusion, the RLK diversity observed in various plant genomes, in terms of domain combinations, did not occur in a single step over the course of plant evolution but instead took place gradually. Clearly, each new sequenced genome, particularly at the basal branches of streptophyte species, has made it possible to describe an increasingly precise history of the origin and evolution of these receptors. Although the picture is still only partial—especially with a lack of data on charophytes—the emergence of certain subfamilies in *C. braunii*, for example, led to the hypothesis that the high morphological complexity of *Chara* could result from the advent and/or expansion of certain gene subfamilies, including LysM, malectin-like (CrRLK1L), and malectin-like LRR RLKs (Figure 4) (106). Similarly, the appearance of RLK subfamilies that are found throughout land plants but missing in charophytes (WAK, L-LEC, CRINKLY4, and S-domain) is correlated with the diversification of both the developmental and defense-signaling mechanisms necessary for out-of-water adaptation (~450 Mya). The appearance of the LRR, C-LEC, and Pro-rich/extensin subfamilies in basal charophytes is still a mystery. How did they emerge? What was their function? Some tracks are beginning to appear, but the way is still long (158). Gaining in-depth understanding of the evolutionary history of these large gene families is a gargantuan task because many events, such as ancestral and recent lineage-specific duplications or domain rearrangements (e.g., fusions or exchanges), must be considered, and the

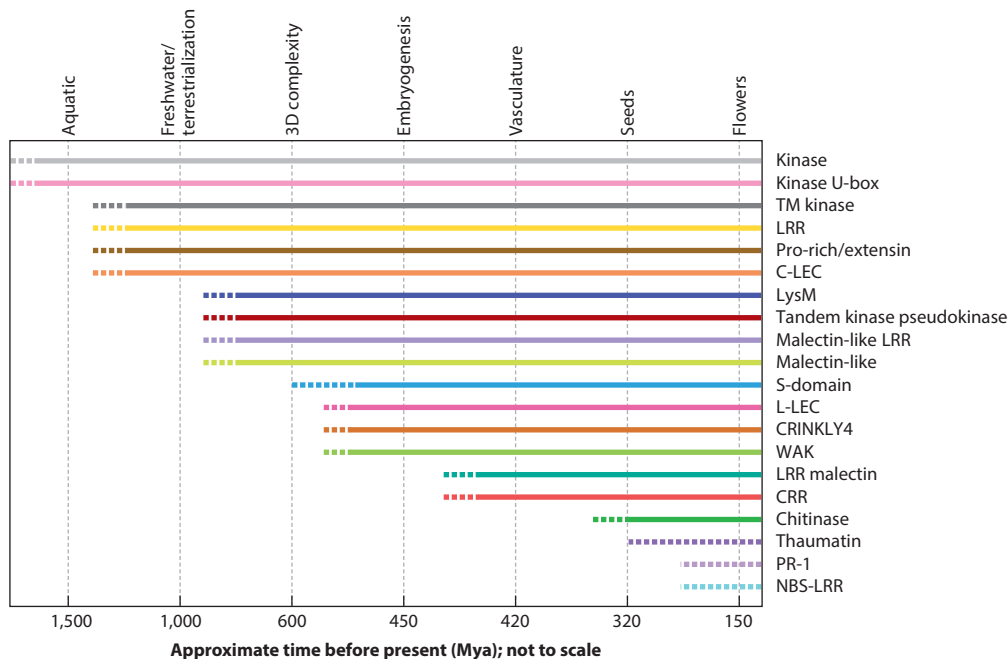


Figure 4

Relationships between the major stages of plant evolution and appearances of new structural organizations of domains. Dotted lines indicate that the exact starting point is unknown. Dates are from References 80 and 120. Abbreviations: C-LEC, calcium-dependent lectin; CRR, cysteine-rich repeat; L-LEC, L-type lectin; LRR, leucine-rich repeat; LysM, lysin motif; NBS, nucleotide-binding site; PR-1, pathogenesis-related protein 1; TM, transmembrane; WAK, wall-associated kinase.

evolutionary dynamics of ECDs and kinase domains must be studied in parallel. This huge work has been undertaken recently for the CRR kinase subfamily (147) and should open the path to new extensive analyses on other subfamilies.

SUMMARY POINTS

1. Plant receptor-like kinases (RLKs) are structurally related to animal receptor tyrosine kinase.
2. Some extracellular domains found in plants are unfamiliar in animals.
3. Domains contained in extracellular regions of plant RLKs are highly diversified.
4. RLKs have a monophyletic origin and are close relatives to animal Pelle and interleukin-1 receptor-associated kinases.
5. RLK diversity, in terms of domain combinations, did not occur in a single step over the course of plant evolution but took place gradually.

NOTE ADDED IN PROOF

While this review was in press, a new receptor configuration was described by Wu, Pei, and coauthors (159). This receptor is the first known cell-surface hydrogen peroxide (HIP) sensor in plants.

The HP-INDUCED Ca²⁺ INCREASES 1 (HPCA1) receptor contains LRRs and a unique domain named the HP domain in its extracellular domain. This HP domain contains two cysteine pairs that are oxidized by HP to activate HPCA1.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was partly supported by the CGIAR Research Program on Rice (CRP RICE) (to A.D.) and by a PhD fellowship of SupAgro and CIRAD (to C.G.). Computational resources were provided by the South Green bioinformatics platform.

LITERATURE CITED

1. Ahlquist RP. 1948. A study of the adrenotropic receptors. *Am. J. Physiol.* 153:586–600
2. Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
3. Ardourel M, Demont N, Debelle F, Maillat F, de Billy F, et al. 1994. *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. *Plant Cell* 6:1357–74
4. Azevedo C, Santos-Rosa MJ, Shirasu K. 2001. The U-box protein family in plants. *Trends Plant Sci.* 6:354–58
5. Bacete L, Mérida H, Miedes E, Molina A. 2018. Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *Plant J.* 93(4):614–36
6. Baek D, Kim MC, Kumar D, Park B, Cheong MS, et al. 2019. AtPR5K2, a PR5-like receptor kinase, modulates plant responses to drought stress by phosphorylating protein phosphatase 2Cs. *Front. Plant Sci.* 10:1146
7. Becraft PW, Stinard PS, McCarty DR. 1996. CRINKLY4: a TNFR-like receptor kinase involved in maize epidermal differentiation. *Science* 273:1406–9
8. Bellande K, Bono JJ, Savelli B, Jamet E, Canut H. 2017. Plant lectins and lectin receptor-like kinases: How do they sense the outside? *Int. J. Mol. Sci.* 18:E1164
9. Bleecker AB. 1991. Genetic analysis of ethylene responses in *Arabidopsis thaliana*. *Symp. Soc. Exp. Biol.* 45:149–58
10. Boisson-Dernier A, Kessler SA, Grossniklaus U. 2011. The walls have ears: the role of plant CrRLK1Ls in sensing and transducing extracellular signals. *J. Exp. Bot.* 62:1581–91
11. Bonner TI, Oppermann H, Seeburg P, Kerby SB, Gunnell MA, et al. 1986. The complete coding sequence of the human raf oncogene and the corresponding structure of the c-raf-1 gene. *Nucleic Acids Res.* 14:1009–15
12. Borassi C, Sede AR, Mecchia MA, Salgado Salter JD, Marzol E, et al. 2016. An update on cell surface proteins containing extensin-motifs. *J. Exp. Bot.* 67(2):477–87
13. Bourdais G, Burdiak P, Gauthier A, Nitsch L, Salojärvi J, et al. 2015. Large-scale phenomics identifies primary and fine-tuning roles for CRKs in responses related to oxidative stress. *PLoS Genet.* 11(7):e1005373
14. Bourne Y, Abergel C, Cambillau C, Frey M, Rouge P, Fontecilla-Camps JC. 1990. X-ray crystal structure determination and refinement at 1.9 Å resolution of isolectin I from the seeds of *Lathyrus ochrus*. *J. Mol. Biol.* 214:571–84
15. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, et al. 2017. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171:287–304.e15

16. Brown GD, Willment JA, Whitehead L. 2018. C-type lectins in immunity and homeostasis. *Nat. Rev. Immunol.* 18(6):374–89
17. Brueggeman R, Druka A, Nirmala J, Cavaleir T, Drader T, et al. 2008. The stem rust resistance gene *Rpg5* encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains. *PNAS* 105:14970–75
18. Brueggeman R, Rostoks N, Kudrna D, Kilian A, Han F, et al. 2002. The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *PNAS* 99:9328–33
19. Buendia 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 Sci.* 9:1531
20. Buist G, Steen A, Kok J, Kuipers OP. 2008. LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.* 68:838–47
21. Chang C, Schaller GE, Patterson SE, Kwok SF, Meyerowitz EM, Bleecker AB. 1992. The *TMK1* gene from *Arabidopsis* codes for a protein with structural and biochemical characteristics of a receptor protein kinase. *Plant Cell* 4:1263–71
22. Chen Z. 2001. A superfamily of proteins with novel cysteine-rich repeats. *Plant Physiol.* 126:473–76
23. Cheng S, Xian W, Fu Y, Marin B, Keller J, et al. 2019. Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell* 179(5):1057–67
24. Choi J, Tanaka K, Cao Y, Qi Y, Qiu J, et al. 2014. Identification of a plant receptor for extracellular ATP. *Science* 343(6168):290–94
25. Clark SE, Williams RW, Meyerowitz EM. 1997. The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* 89:575–85
26. Cook DE, Mesarich CH, Thomma BPHJ. 2015. Understanding plant immunity as a surveillance system to detect invasion. *Annu. Rev. Phytopathol.* 53:541–63
27. Couto D, Zipfel C. 2016. Regulation of pattern recognition receptor signalling in plants. *Nat. Rev. Immunol.* 16(9):537–52
28. Czyzewicz N, Nikonorova N, Meyer MR, Sandal P, Shah S, et al. 2016. The growing story of (ARA-BIDOPSIS) CRINKLY 4. *J. Exp. Bot.* 67(16):4835–47
29. Dangl JL, Preuss D, Schroeder JI. 1995. Talking through walls: signaling in plant development. *Cell* 83:1071–77
30. Dardick C, Schwessinger B, Ronald P. 2012. Non-arginine-aspartate (non-RD) kinases are associated with innate immune receptors that recognize conserved microbial signatures. *Curr. Opin. Plant Biol.* 15(4):358–66
31. Day EK, Sosale NG, Lazzara MJ. 2016. Cell signaling regulation by protein phosphorylation: a multi-variate, heterogeneous, and context-dependent process. *Curr. Opin. Biotechnol.* 40:185–92
32. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, et al. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 34:W362–65
33. Deeken R, Kaldenhoff R. 1997. Light-repressible receptor protein kinase: a novel photo-regulated gene from *Arabidopsis thaliana*. *Planta* 202:479–86
34. Delaux PM, Radhakrishnan GV, Jayaraman D, Cheem J, Malbreil M, et al. 2015. Algal ancestor of land plants was preadapted for symbiosis. *PNAS* 112:13390–95
35. Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, et al. 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *PNAS* 103:11647–52
36. Dievart A, Gilbert N, Droc G, Attard A, Gourgues M, et al. 2011. Leucine-rich repeat receptor kinases are sporadically distributed in eukaryotic genomes. *BMC Evol. Biol.* 11:367
37. Du S, Qu LJ, Xiao J. 2018. Crystal structures of the extracellular domains of the CrRLK1L receptor-like kinases ANXUR1 and ANXUR2. *Protein Sci.* 27(4):886–92
38. Dufayard JF, Bettembourg M, Fischer I, Droc G, Guiderdoni E, et al. 2017. New insights on leucine-rich repeats receptor-like kinase orthologous relationships in angiosperms. *Front. Plant Sci.* 8:381
39. Eбина Y, Ellis L, Jarnagin K, Edery M, Graf L, et al. 1985. The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling. *Cell* 40:747–58

40. Fan J, Bai P, Ning Y, Wang J, Shi X, et al. 2018. The monocot-specific receptor-like kinase SDS2 controls cell death and immunity in rice. *Cell Host Microbe* 23(4):498–510.E5
41. Feuillet C, Reuzeau C, Kjellbom P, Keller B. 1998. Molecular characterization of a new type of receptor-like kinase (wlrk) gene family in wheat. *Plant Mol. Biol.* 37:943–53
42. Fischer I, Dievart A, Droc G, Dufayard JF, Chantret N. 2016. Evolutionary dynamics of the leucine-rich repeat receptor-like kinase (LRR-RLK) subfamily in angiosperms. *Plant Physiol.* 170:1595–610
43. Ge Z, Dresselhaus T, Qu LJ. 2019. How CrRLK1L receptor complexes perceive RALF signals. *Trends Plant Sci.* 24(11):978–81
44. Goldberg JM, Manning G, Liu A, Fey P, Pilcher KE, et al. 2006. The dictyostelium kinome—analysis of the protein kinases from a simple model organism. *PLOS Genet.* 2:e38
45. Guo Y, Peng D, Zhou J, Lin S, Wang C, et al. 2019. iEKPd 2.0: an update with rich annotations for eukaryotic protein kinases, protein phosphatases and proteins containing phosphoprotein-binding domains. *Nucleic Acids Res.* 47:D344–50
46. Gust AA, Felix G. 2014. Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. *Curr. Opin. Plant Biol.* 21:104–11
47. Hanks SK, Hunter T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* 9:576–96
48. Hanks SK, Quinn AM, Hunter T. 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241:42–52
49. Hardie DG. 1999. Plant protein serine/threonine kinases: classification and functions. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50:97–131
50. Hashimoto C, Hudson KL, Anderson KV. 1988. The *Toll* gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell* 52:269–79
51. He Y, Zhou J, Shan L, Meng X. 2018. Plant cell surface receptor-mediated signaling—a common theme amid diversity. *J. Cell Sci.* 131:jcs209353
52. He ZH, Cheeseman I, He D, Kohorn BD. 1999. A cluster of five cell wall-associated receptor kinase genes, *Wak1–5*, are expressed in specific organs of *Arabidopsis*. *Plant Mol. Biol.* 39:1189–96
53. He ZH, Fujiki M, Kohorn BD. 1996. A cell wall-associated, receptor-like protein kinase. *J. Biol. Chem.* 271:19789–93
54. Hedges SB. 2002. The origin and evolution of model organisms. *Nat. Rev. Genet.* 3:838–49
55. Herve C, Dabos P, Galaud JP, Rouge P, Lescure B. 1996. Characterization of an *Arabidopsis thaliana* gene that defines a new class of putative plant receptor kinases with an extracellular lectin-like domain. *J. Mol. Biol.* 258:778–88
56. Hirayama T, Oka A. 1992. Novel protein kinase of *Arabidopsis thaliana* (APK1) that phosphorylates tyrosine, serine and threonine. *Plant Mol. Biol.* 20:653–62
57. Hohmann U, Lau K, Hothorn M. 2017. The structural basis of ligand perception and signal activation by receptor kinases. *Annu. Rev. Plant Biol.* 68:109–37
58. Hohmann U, Santiago J, Nicolet J, Olsson V, Spiga FM, et al. 2018. Mechanistic basis for the activation of plant membrane receptor kinases by SERK-family coreceptors. *PNAS* 115(13):3488–93
59. Hok S, Danchin EG, Allasia V, Panabieres F, Attard A, Keller H. 2011. An *Arabidopsis* (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. *Plant Cell Environ.* 34:1944–57
60. Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, et al. 2014. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* 5:3978
61. Kerk D, Bulgrien J, Smith DW, Gribskov M. 2003. *Arabidopsis* proteins containing similarity to the universal stress protein domain of bacteria. *Plant Physiol.* 131:1209–19
62. Kim YS, Lee JH, Yoon GM, Cho HS, Park SW, et al. 2000. CHRK1, a chitinase-related receptor-like kinase in tobacco. *Plant Physiol.* 123:905–15
63. Klymiuk V, Yaniv E, Huang L, Raats D, Fatiukha A, et al. 2018. Cloning of the wheat *Yr15* resistance gene sheds light on the plant tandem kinase-pseudokinase family. *Nat. Commun.* 9:3735
64. Kohorn BD. 2016. Cell wall-associated kinases and pectin perception. *J. Exp. Bot.* 67(2):489–94

65. Kohorn BD, Lane S, Smith TA. 1992. An *Arabidopsis* serine/threonine kinase homologue with an epidermal growth factor repeat selected in yeast for its specificity for a thylakoid membrane protein. *PNAS* 89:10989–92
66. Krattinger SG, Keller B. 2016. Molecular genetics and evolution of disease resistance in cereals. *New Phytol.* 212(2):320–32
67. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305:567–80
68. Kroj T, Chanclud E, Michel-Romiti C, Grand X, Morel JB. 2016. Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *New Phytol.* 210(2):618–26
69. Krupa A, Abhinandan KR, Srinivasan N. 2004. KinG: a database of protein kinases in genomes. *Nucleic Acids Res.* 32:D153–55
70. Krupa A, Anamika, Srinivasan N. 2006. Genome-wide comparative analyses of domain organisation of repertoires of protein kinases of *Arabidopsis thaliana* and *Oryza sativa*. *Gene* 380:1–13
71. Kutschera A, Dawid C, Gisch N, Schmid C, Raasch L, et al. 2019. Bacterial medium-chain 3-hydroxy fatty acid metabolites trigger immunity in *Arabidopsis* plants. *Science* 364(6436):178–81
72. Kwon A, Scott S, Taujale R, Yeung W, Kochut KJ, et al. 2019. Tracing the origin and evolution of pseudokinases across the tree of life. *Sci. Signal.* 12(578):eaav3810
73. Labbé J, Muchero W, Czarnecki O, Wang J, Wang X, et al. 2019. Mediation of plant-mycorrhizal interaction by a lectin receptor-like kinase. *Nat. Plants* 5(7):676–80
74. Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, et al. 2018. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* 93:515–33
75. Lee DS, Kim YC, Kwon SJ, Ryu CM, Park OK. 2017. The *Arabidopsis* cysteine-rich receptor-like kinase CRK36 regulates immunity through interaction with the cytoplasmic kinase BIK1. *Front. Plant Sci.* 8:1856
76. Lee JH, Takei K, Sakakibara H, Sun Cho H, Kim DM, et al. 2003. CHRK1, a chitinase related receptor-like kinase, plays a role in plant development and cytokinin homeostasis in tobacco. *Plant Mol. Biol.* 53(6):877–90
77. Lehti-Shiu MD, Shiu SH. 2012. Diversity, classification and function of the plant protein kinase superfamily. *Philos. Trans. R. Soc. B* 367:2619–39
78. Lehti-Shiu MD, Zou C, Hanada K, Shiu SH. 2009. Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* 150:12–26
79. Lehti-Shiu MD, Zou C, Shiu SH. 2012. Origin, diversity, expansion history, and functional evolution of the plant receptor-like kinase/pelle family. In *Receptor-Like Kinases in Plants: Signaling and Communication in Plants*, Vol. 13, ed. F Tax, B Kemmerling, pp. 1–22. Berlin: Springer
80. Leliaert F, Verbruggen H, Zechman FW. 2011. Into the deep: new discoveries at the base of the green plant phylogeny. *Bioessays* 33:683–92
81. Lemieux C, Otis C, Turmel M. 2007. A clade uniting the green algae *Mesostigma viride* and *Chlorokybus atmophyticus* represents the deepest branch of the Streptophyta in chloroplast genome-based phylogenies. *BMC Biol.* 5:2
82. Li B, Ferreira MA, Huang M, Camargos LF, Yu X, et al. 2019. The receptor-like kinase NIK1 targets FLS2/BAK1 immune complex and inversely modulates antiviral and antibacterial immunity. *Nat. Commun.* 10(1):4996
83. Li J, Chory J. 1997. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* 90:929–38
84. Li W, Liu Y, Wang J, He M, Zhou X, et al. 2016. The durably resistant rice cultivar Digu activates defence gene expression before the full maturation of *Magnaporthe oryzae* appressorium. *Mol. Plant Pathol.* 17(3):354–68
85. Liang X, Zhou JM. 2018. Receptor-like cytoplasmic kinases: central players in plant receptor kinase-mediated signaling. *Annu. Rev. Plant Biol.* 69:267–99
86. Liebrand TW, van den Burg HA, Joosten MH. 2014. Two for all: receptor-associated kinases SOBIR1 and BAK1. *Trends Plant. Sci.* 19(2):123–32

87. Liener IE. 1964. Seed hemagglutinins. *Econ. Bot.* 18:27–33
88. Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R. 2003. LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* 302:630–33
89. Limpens E, van Zeijl A, Geurts R. 2015. Lipochitooligosaccharides modulate plant host immunity to enable endosymbioses. *Annu. Rev. Phytopathol.* 53:311–34
90. Liu PL, Du L, Huang Y, Gao SM, Yu M. 2017. Origin and diversification of leucine-rich repeat receptor-like protein kinase (*LRR-RLK*) genes in plants. *BMC Evol. Biol.* 17:47
91. Loh YT, Martin GB. 1995. The disease-resistance gene *PTO* and the fenthion-sensitivity gene *FEN* encode closely related functional protein kinases. *PNAS* 92:4181–84
92. Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, et al. 2003. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* 425:637–40
93. Maehle A-H. 2004. “Receptive substances”: John Newport Langley (1852–1925) and his path to a receptor theory of drug action. *Med. Hist.* 48:153–74
94. Maehle A-H. 2009. A binding question: the evolution of the receptor concept. *Endeavour* 33:135–40
95. Manning G, Plowman GD, Hunter T, Sudarsanam S. 2002. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* 27:514–20
96. Manning G, Reiner DS, Lauwaet T, Dacre M, Smith A, et al. 2011. The minimal kinome of *Giardia lamblia* illuminates early kinase evolution and unique parasite biology. *Genome Biol.* 12:R66
97. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. 2002. The protein kinase complement of the human genome. *Science* 298:1912–34
98. Martin DM, Miranda-Saavedra D, Barton GJ. 2009. Kinomer v. 1.0: a database of systematically classified eukaryotic protein kinases. *Nucleic Acids Res.* 37:D244–50
99. Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganai MW, et al. 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262:1432–36
100. Martinez M. 2016. Computational tools for genomic studies in plants. *Curr. Genom.* 17:509–14
101. McCarty DR, Chory J. 2000. Conservation and innovation in plant signaling pathways. *Cell* 103:201–9
102. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, et al. 2015. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 43:D213–21
103. Moussu S, Augustin S, Roman AO, Broyart C, Santiago J. 2018. Crystal structures of two tandem malectin-like receptor kinases involved in plant reproduction. *Acta Crystallogr. D Struct. Biol.* 74(Pt 7):671–80
104. Nasrallah JB, Kao TH, Chen CH, Goldberg ML, Nasrallah ME. 1987. Amino-acid sequence of glycoproteins encoded by three alleles of the *S* locus of *Brassica oleracea*. *Nature* 326:617–19
105. Nikonorova N, Vu LD, Czyzewicz N, Gevaert K, De Smet I. 2015. A phylogenetic approach to study the origin and evolution of the CRINKLY4 family. *Front. Plant Sci.* 6:880
106. Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, et al. 2018. The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell* 174:448–64.E24
107. Nissen KS, Willats WGT, Malinovsky FG. 2016. Understanding CrRLK1L function: cell walls and growth control. *Trends Plant Sci.* 21(6):516–27
108. Pan H, Stonoha-Arther C, Wang D. 2018. Medicago plants control nodulation by regulating proteolysis of the receptor-like kinase DMI2. *Plant Physiol.* 177(2):792–802
109. Pan J, Li Z, Wang Q, Yang L, Yao F, Liu W. 2020. An S-domain receptor-like kinase, OsESG1, regulates early crown root development and drought resistance in rice. *Plant Sci.* 290:110318
110. Park J, Kim TH, Takahashi Y, Schwab R, Dressano K, et al. 2019. Chemical genetic identification of a lectin receptor kinase that transduces immune responses and interferes with abscisic acid signaling. *Plant J.* 98(3):492–510
111. Pelagio-Flores R, Muñoz-Parra E, Barrera-Ortiz S, Ortiz-Castro R, Saenz-Mata J, et al. 2019. The cysteine-rich receptor-like protein kinase CRK28 modulates *Arabidopsis* growth and development and influences abscisic acid responses. *Planta* 251(1):2
112. Pruitt RE, Hulskamp M, Kopeczak SD, Ploense SE, Schneitz K. 1993. Molecular genetics of cell interactions in *Arabidopsis*. *Dev. Suppl.* 1993:77–84

113. Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, et al. 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425:585–92
114. Rapp UR, Goldsborough MD, Mark GE, Bonner TI, Groffen J, et al. 1983. Structure and biological activity of v-raf, a unique oncogene transduced by a retrovirus. *PNAS* 80:4218–22
115. Sarris PF, Cevik V, Dagdas G, Jones JD, Krasileva KV. 2016. Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biol.* 14:8
116. Sasaki G, Katoh K, Hirose N, Suga H, Kuma K, et al. 2007. Multiple receptor-like kinase cDNAs from liverwort *Marchantia polymorpha* and two charophycean green algae, *Closterium ehrenbergii* and *Nitella axillaris*: extensive gene duplications and gene shufflings in the early evolution of streptophytes. *Gene* 401:135–44
117. Schellenberger R, Touchard M, Clément C, Baillieul F, Cordelier S, et al. 2019. Apoplastic invasion patterns triggering plant immunity: plasma membrane sensing at the frontline. *Mol. Plant Pathol.* 20(11):1602–16
118. Schlessinger J. 2014. Receptor tyrosine kinases: legacy of the first two decades. *Cold Spring Harb. Perspect. Biol.* 6:a008912
119. Schulze-Muth P, Irmeler S, Schröder G, Schröder J. 1996. Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*). cDNA, gene, intramolecular autophosphorylation, and identification of a threonine important for auto- and substrate phosphorylation. *J. Biol. Chem.* 271:26684–89
120. Sessa EB, Banks JA, Barker MS, Der JP, Duffy AM, et al. 2014. Between two fern genomes. *GigaScience* 3:15
121. Shiu SH, Bleecker AB. 2001. Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *PNAS* 98:10763–68
122. Shiu SH, Bleecker AB. 2003. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiol.* 132:530–43
123. Shiu SH, Karlowski WM, Pan R, Tzeng YH, Mayer KF, Li WH. 2004. Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. *Plant Cell* 16:1220–34
124. Shumayla SS, Sharma S, Pandey AK, Singh K, Upadhyay SK. 2016. Molecular characterization and global expression analysis of lectin receptor kinases in bread wheat (*Triticum aestivum*). *PLOS ONE* 11:e0153925
125. Silva NF, Goring DR. 2002. The proline-rich, extensin-like receptor kinase-1 (PERK1) gene is rapidly induced by wounding. *Plant Mol. Biol.* 50:667–85
126. Silverstein AM. 2002. *Paul Ehrlich's Receptor Immunology: The Magnificent Obsession*. San Diego, CA: Academic
127. Simon A, Glockner G, Felder M, Melkonian M, Becker B. 2006. EST analysis of the scaly green flagellate *Mesostigma viride* (Streptophyta): implications for the evolution of green plants (Viridiplantae). *BMC Plant Biol.* 6:2
128. Smakowska-Luzan E, Mott GA, Parys K, Stegmann M, Howton TC, et al. 2018. An extracellular network of *Arabidopsis* leucine-rich repeat receptor kinases. *Nature* 553(7688):342–46
129. Snyders S, Kohorn BD. 1999. TAKs, thylakoid membrane protein kinases associated with energy transduction. *J. Biol. Chem.* 274:9137–40
130. Soanes DM, Talbot NJ. 2010. Comparative genome analysis reveals an absence of leucine-rich repeat pattern-recognition receptor proteins in the kingdom Fungi. *PLOS ONE* 5:e12725
131. Song WY, Han Z, Wang J, Lin G, Chai J. 2017. Structural insights into ligand recognition and activation of plant receptor kinases. *Curr. Opin. Struct. Biol.* 43:18–27
132. Song WY, Wang GL, Chen LL, Kim HS, Pi LY, et al. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 270:1804–6
133. Stegmann M, Monaghan J, Smakowska-Luzan E, Rovenich H, Lehner A, et al. 2017. The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* 355(6322):287–89
134. Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB. 1991. Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. *PNAS* 88:8816–20
135. Stillmark H. 1889. Uber Ricin. *Arch. Pharmakol. Inst. Dorpat.* 3:59

136. Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, et al. 2002. A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417:959–62
137. Takahashi T, Mu JH, Gasch A, Chua NH. 1998. Identification by PCR of receptor-like protein kinases from *Arabidopsis* flowers. *Plant Mol. Biol.* 37:587–96
138. Tamborski J, Krasileva KV. 2020. Evolution of plant NLRs: from natural history to precise modifications. *Annu. Rev. Plant Biol.* 71:355–78
139. Tang D, Wang G, Zhou JM. 2017. Receptor kinases in plant-pathogen interactions: more than pattern recognition. *Plant Cell* 29:618–37
140. Teixeira PJ, Costa GG, Fiorin GL, Pereira GA, Mondego JM. 2013. Novel receptor-like kinases in cacao contain PR-1 extracellular domains. *Mol. Plant Pathol.* 14:602–9
141. Tobias CM, Howlett B, Nasrallah JB. 1992. An *Arabidopsis thaliana* gene with sequence similarity to the S-locus receptor kinase of *Brassica oleracea*: sequence and expression. *Plant Physiol.* 99:284–90
142. Tordai H, Bányai L, Patthy L. 1999. The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins. *FEBS Lett.* 461:63–67
143. Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, et al. 1996. The *Arabidopsis* ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* 8:735–46
144. Uknes S, Mauch-Mani B, Moyer M, Potter S, Williams S, et al. 1992. Acquired resistance in *Arabidopsis*. *Plant Cell* 4:645–56
145. Ullrich A, Bell JR, Chen EY, Herrera R, Petruzzelli LM, et al. 1985. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature* 313:756–61
146. Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, et al. 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature* 309:418–25
147. Vaattovaara A, Brandt B, Rajaraman S, Safronov O, Veidenberg A, et al. 2019. Mechanistic insights into the evolution of DUF26-containing proteins in land plants. *Commun. Biol.* 2:56
148. Vaid N, Pandey PK, Tuteja N. 2012. Genome-wide analysis of lectin receptor-like kinase family from *Arabidopsis* and rice. *Plant Mol. Biol.* 80:365–88
149. Valon C, Smalle J, Goodman HM, Giraudat J. 1993. Characterization of an *Arabidopsis thaliana* gene (TMKL1) encoding a putative transmembrane protein with an unusual kinase-like domain. *Plant Mol. Biol.* 23:415–21
150. van der Wel H, Loeve K. 1972. Isolation and characterization of thaumatin I and II, the sweet-tasting proteins from *Thaumatococcus daniellii* Benth. *Eur. J. Biochem.* 31:221–25
151. Vernié T, Camut S, Camps C, Rembliere C, de Carvalho-Niebel F, et al. 2016. PUB1 interacts with the receptor kinase DMI2 and negatively regulates rhizobial and arbuscular mycorrhizal symbioses through its ubiquitination activity in *Medicago truncatula*. *Plant Physiol.* 170(4):2312–24
152. Vigneron N, Radhakrishnan GV, Delaux PM. 2018. What have we learnt from studying the evolution of the arbuscular mycorrhizal symbiosis? *Curr. Opin. Plant Biol.* 44:49–56
153. Walker JC. 1993. Receptor-like protein kinase genes of *Arabidopsis thaliana*. *Plant J.* 3:451–56
154. Walker JC, Zhang R. 1990. Relationship of a putative receptor protein kinase from maize to the S-locus glycoproteins of *Brassica*. *Nature* 345:743–46
155. Wang C, Huang X, Li Q, Zhang Y, Li JL, Mou Z. 2019. Extracellular pyridine nucleotides trigger plant systemic immunity through a lectin receptor kinase/BAK1 complex. *Nat. Commun.* 10(1):4810
156. Wang X, Zafian P, Choudhary M, Lawton M. 1996. The PR5K receptor protein kinase from *Arabidopsis thaliana* is structurally related to a family of plant defense proteins. *PNAS* 93:2598–602
157. Wang Y, Bouwmeester K. 2017. L-type lectin receptor kinases: new forces in plant immunity. *PLOS Pathog.* 13(8):e1006433
158. Whitewoods C, Cammarata J, Nemeček V, Sang S, Crook A, et al. 2018. *CLAVATA* was a genetic novelty for the morphological innovation of 3D growth in land plants. *Curr. Biol.* 28:2365–76
159. Wu F, Chi Y, Jiang Z, Xu Y, Xie L, et al. 2020. Hydrogen peroxide sensor HPCA1 is an LRR receptor kinase in *Arabidopsis*. *Nature* 578:577–81

160. Xi L, Wu XN, Gilbert M, Schulze WX. 2019. Classification and interactions of LRR receptors and co-receptors within the Arabidopsis plasma membrane—an overview. *Front. Plant Sci.* 10:472
161. Xiao Y, Stegmann M, Han Z, DeFalco TA, Parys K, et al. 2019. Mechanisms of RALF peptide perception by a heterotypic receptor complex. *Nature* 572(7768):270–74
162. Xing S, Li M, Liu P. 2013. Evolution of S-domain receptor-like kinases in land plants and origination of S-locus receptor kinases in Brassicaceae. *BMC Evol. Biol.* 13:69
163. Yang Y, Labbe J, Muchero W, Yang X, Jawdy SS, et al. 2016. Genome-wide analysis of lectin receptor-like kinases in *Populus*. *BMC Genom.* 17:699
164. Yeh YH, Panzeri D, Kadota Y, Huang YC, Huang PY, et al. 2016. The Arabidopsis malectin-like/LRR-RLK IOS1 is critical for BAK1-dependent and BAK1-independent pattern-triggered immunity. *Plant Cell* 28(7):1701–21
165. Zebisch A, Troppmair J. 2006. Back to the roots: the remarkable RAF oncogene story. *Cell Mol. Life Sci.* 63:1314–30
166. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, et al. 2016. iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* 9:1667–70
167. Zulawski M, Schulze G, Braginets R, Hartmann S, Schulze WX. 2014. The Arabidopsis Kinome: phylogeny and evolutionary insights into functional diversification. *BMC Genom.* 15:548