Frans Tax
Birgit Kemmerling   *Editors*

Receptor-like Kinases in Plants

From Development to Defense

Springer

Signaling and Communication in Plants

Frans Tax • Birgit Kemmerling
Editors

Receptor-like Kinases in Plants

From Development to Defense

Editors
Frans Tax
University of Arizona
Department of Molecular and
Cellular Biology
1007 E. Lowell
Tucson, AZ 85721
USA
fetax@email.arizona.edu

Birgit Kemmerling
University of Tuebingen
Department of Plant
Biochemistry
Auf der Morgenstelle 5
72076 Tuebingen
Germany
birgit.kemmerling@zmbp.uni-tuebingen.de

Printed on acid-free paper

Preface

In 1991, John Walker and Ren Zhang reported the identification of the first plant receptor protein kinase gene (1). Based on its predicted structure of an extracellular domain related to the Self-incompatibility (S)-locus secreted glycoprotein, and a cytoplasmic serine/threonine kinase domain, Walker and Zhang predicted that the discovery of these novel plant proteins "provides a unique opportunity to gain fresh insights into signal transduction in higher plants".

Twenty years after Walker and Zhang's initial findings, it is difficult to imagine plant biology without receptor kinases coming to mind. The next decade saw the emergence of additional receptor kinases through forward genetic screens and through molecular biology. As alluded to above, the S-locus in Brassica also encoded receptor kinases with extracellular domains related to the S-locus glyco-proteins. Developmental functions for various receptor kinases included epidermal differentiation for the CRINKLY-4 gene in maize, morphogenesis for the ERECTA gene in Arabidopsis, and maintenance of stem cells in the shoot apical meristem for the CLAVATA-1 gene in Arabidopsis. A role for receptor kinases in recognition of pathogens was first revealed by the identification of Xa21 in rice. A big surprise came with the finding that the BRI1 receptor kinase was the receptor for the plant steroid hormone brassinosteroids.

The emerging genome sequence of Arabidopsis was also uncovering hundreds of receptor kinases, ultimately more than 600. In the report of the NSF-Sponsored Workshop: "New Directions in Plant Biological Research" in April of 1999 (http://www.arabidopsis.org/carnegie_rep.html), the authors asked: "What are the roles of the hundreds of these proteins? Their existence implies a massive network of cell–cell and environment–plant communication, via a series of ligands yet to be discovered. Understanding this network will give us an entirely new view of plant development, environmental response, and organismal integration." Analysis of other genomes, including rice, which has more than 900 receptor kinases, indicates that the large number of receptor kinases in Arabidopsis was not an anomaly.

This book focuses on the momentum created within the plant biology community since Walker and Zhang's initial discovery. Thanks to a combination of collaborative "omics" projects, as well as the deep research efforts of many labs, portions of this "massive network" are emerging. This book opens with a view of the evolution and conservation of receptor kinases in plants, focusing on the rapid expansion of this gene family. After the first chapter, the following seven chapters update the known functions of receptor kinases in various biological contexts, extending the initial discoveries mentioned above. The second half of the book focuses on the diverse ligands, signaling mechanisms, and regulation of receptor kinases. The authors of all of these chapters reveal the amazing results from the past 20 years, and hint at the discoveries that may come in the next 20 years.

Tucson, USA   *Frans Tax*
Tübingen, Germany   *Birgit Kemmerling*

Contents

Origin, Diversity, Expansion History, and Functional Evolution of the Plant Receptor-Like Kinase/*Pelle* Family

Melissa D. Lehti-Shiu, Cheng Zou, and Shin-Han Shiu

Abstract The RLK/*Pelle* gene family is one of the largest gene families in plants with several hundred to more than a thousand members, but only a few family members exist in animals. This unbalanced distribution indicates a rather dramatic expansion of this gene family in land plants. In this chapter we review what is known about the RLK/*Pelle* family's origin in eukaryotes, its domain content evolution, expansion patterns across plant and animal species, and the duplication mechanisms that contribute to its expansion. We conclude by summarizing current knowledge of plant RLK/*Pelle* functions for a discussion on the relative importance of neutral evolution and natural selection as the driving forces behind continuous expansion and innovation in this gene family.

1 Introduction

In 1990, John Walker and Ren Zhang reported the cloning of a maize protein kinase resembling animal receptor tyrosine kinases (Walker and Zhang 1990). This maize kinase, which contains a putative extracellular domain (ECD) delineated by a signal sequence and a hydrophobic transmembrane region, represents the prototypical Receptor-Like Kinase (RLK). In the following two decades, extensive genetic and phenotypic studies revealed diverse roles of RLKs, ranging from control of development to stress responses (Walker 1994; Braun and Walker 1996; Torii 2000; Shiu and Bleecker 2001b; Lease and Walker 2006; Morillo and Tax 2006). In the late 1990s, as expressed sequence tags and genomic sequence accumulated, more and more RLKs were found in the model plant *Arabidopsis thaliana*. A global computational analysis of *A. thaliana* kinases then established that RLKs were one of the largest gene families in plants (Shiu and Bleecker 2001a). Interestingly, RLKs were

M.D. Lehti-Shiu • C. Zou • S.-H. Shiu (✉)
Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA
e-mail: lehtishi@msu.edu; czou@caas.net.cn; shius@msu.edu

found to be most closely related to *Drosophila melanogaster Pelle* (Belvin and Anderson 1996) and mammalian Interleukin Receptor-Associated Kinases (IRAKs) (Cao et al. 1996; Flannery and Bowie 2010), which comprise a very small family of cytoplasmic kinases without ECDs or membrane spanning regions. Similar to *Pelle* and IRAKs, some plant members of the RLK/*Pelle* family also lack ECDs and transmembrane regions and are referred to as Receptor-Like Cytoplasmic Kinases (RLCKs).The phylogenetic relationships between RLKs and *Pelle*/IRAKs indicate that they likely are orthologs derived from an ancestral kinase present in the common ancestor of plants and animals. Thus, these related kinases are collectively classified as members of the RLK/*Pelle* family (Shiu and Bleecker 2001a).

In this chapter, we discuss the current knowledge of the evolutionary history of the RLK/*Pelle* family with a focus on the following areas. Our first focus is on the origin of the RLK/*Pelle* family, particularly on the relationship between the kinase domains of plant RLK/*Pelle*s and other kinase families. The second focus concerns

Table 1 Protein domains found in the extracellular regions of RLK/*Pelles*

Protein domain	Description	Reference
Bulb-type lectin domain (B-lectin)	Binds α-D-mannose	Hester et al. (1995)
Cyclases/Histidine kinases Associated Sensory Extracellular (CHASE)	Found in diverse transmembrane receptors and is predicted to bind low molecular weight ligands	Mougel and Zhulin (2001)
C-type lectin (C-LEC)	Binds carbohydrates	Sharon and Lis (2001)
Domain of Unknown Function26 (DUF26)	Contains four conserved cysteines and implicated in defense response	Miyakawa et al. (2007)
Epidermal Growth Factor (EGF)	Contains calcium binding motif and calcium binding may be necessary for protein–protein interactions	Handford et al. (1990)
Glycerophosphodiester Phosphodiesterase Domain (GDPD)	Found in enzymes that hydrolyze glycerophosphodiesters	Santelli et al. (2004)
Glycoside hydrolase, family 18 catalytic domain	Hydrolyzes chitin oligosaccharides	Perrakis et al. (1994)
Legume lectin B	Carbohydrate binding	Loris et al. (1998)
Lysin Motif (LysM)	Peptidoglycan, chitin-binding	Buist et al. (2008)
Leucine-Rich Repeat (LRR)	Protein–protein interactions	Kobe and Deisenhofer (1994)
Plasminogen/Apple/Nematode protein domain (PAN)	Protein–protein and protein–carbohydrate interactions	Tordai et al. (1999)
S-locus glycoprotein	Secreted proteins involved in self-incompatibility	Hinata et al. (1995)
Pathogenesis-Related Protein-1/Sperm-Coating Glycoprotein domain (SCP)	Found in pathogenesis-related proteins expressed during defense response	Szyperski et al. (1998)
Thaumatin	Antifungal and chitinase activity	Pan et al. (1999)
WAK	Domain found in Wall-associated Kinases, which bind cell wall components	Kohorn (2001)

how the RLK/*Pelle* family has diversified in plants. Plant RLKs with a clear receptor configuration possess a diverse array of extracellular regions implicated in interactions with proteins, polysaccharides, lipids, and other ligands (Shiu and Bleecker 2001b) (see Table 1, Chapters "Receptor Ligands in Development" and "Phosphorylation and RLK signaling"). In particular, we will address two questions regarding RLK diversification: (1) how RLKs acquired their extracellular regions and (2) how frequently these acquisitions have occurred. Third, we will focus on how the RLK/*Pelle* family expanded during plant evolution. The RLK/*Pelle* family is among the largest in plants, but there is only one *D. melanogaster Pelle* kinase and four human IRAKs (Belvin and Anderson 1996; Janssens and Beyaert 2003). Why is there such a large size difference between plant and animal RLK/*Pelle* members? Based on kinase phylogeny, the RLK/*Pelle* family can be subdivided into multiple subfamilies where members in each subfamily possess mostly similar ECDs (Shiu and Bleecker 2001a). How have these subfamilies expanded differentially among plant lineages and what are the implications for plant adaptation and evolution? Finally, we will summarize current knowledge of plant RLK/*Pelle* functions to pinpoint potential sources of selective pressure that drive continuous expansion and innovation in this gene family.

2 Origin of Receptor-Like/*Pelle* Kinases in Eukaryotes

Phylogenetic studies of representative kinase family members indicate that plant RLKs and animal *Pelle* kinases and IRAKs are each others' closest relative and that the RLK/*Pelle* family predates the divergence of the plant and animal lineages (Shiu and Bleecker 2001a). In addition, the kinase domains of RLK/*Pelle*s are more closely related to animal receptor tyrosine kinases and Raf kinases than to any other kinase family (Shiu and Bleecker 2001a). Thus, RLK/*Pelle* and receptor tyrosine kinases likely have a monophyletic origin. Interestingly, aside from G-protein coupled receptors, receptor tyrosine kinases belong to the largest family of transmembrane receptors in animals (Hunter et al. 1992; van der Geer et al. 1994). Taking into consideration that plant RLKs and animal receptor tyrosine kinases have highly similar structural configurations and are likely monophyletic in origin, it is likely that their common ancestor was involved in the perception of extracellular stimuli. However, it is not clear if the ancestral kinase of RLK/*Pelle* and tyrosine kinases was a transmembrane receptor protein or a cytoplamic kinase that interacted with a receptor complex that perceived extracellular signals.

The RLK/*Pelle* family was originally established using plant and metazoan homologs only (Shiu and Bleecker 2001b). Subsequent studies with expanded taxonomic sampling have revealed that there are no clear RLK/*Pelle* homologs in fungi (Shiu and Bleecker 2003). For this review, we have also surveyed multiple nonplant eukaryotic genomes currently available in GenBank. Our findings confirm the absence of RLK/*Pelle* homologs in fungi and indicate that the taxonomic distribution of RLK/*Pelle*s is surprisingly sparse among eukaryotes (Fig. 1).
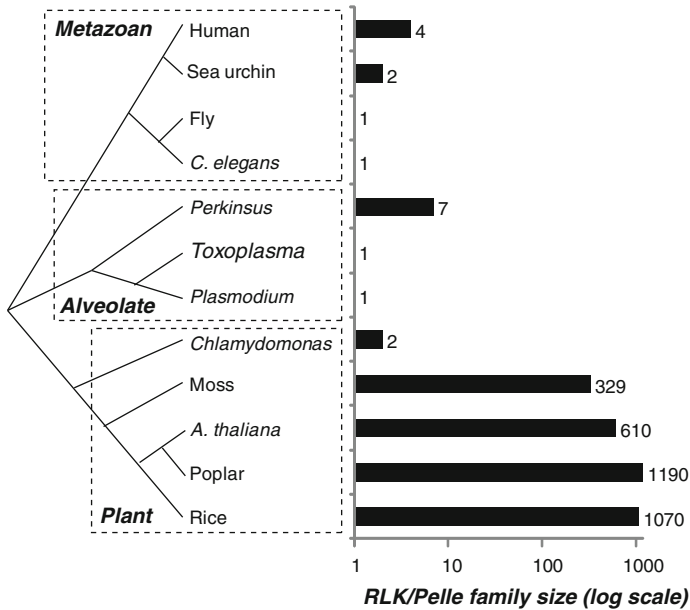
Fig. 1 Phylogenetic relationships between species harboring RLK/*Pelle* homologs and the RLK family size in each species

Although *Pelle* and IRAK related sequences can be found in multiple vertebrate and invertebrate animals, there are no clear homologs in *Monosiga brevicollis*, a unicellular choanoflagellate that is basal to all metazoans (King et al. 2008). In addition, among nonplant and nonmetazoan eukaryotes, RLK/*Pelle* homologs are only found in alveolata species including *Plasmodium*, *Toxoplasma*, and *Perkinsus*. The distribution of RLK/*Pelle* genes among eukaryotic species suggests an ancient origin before the divergence between the plant and animal lineages. The absence of RLK/*Pelle* in most eukaryotic lineages can be explained by either gene losses in multiple lineages or sequence divergence. The difficulty in using gene loss as an explanation is that multiple independent losses would have had to occur to account for the patterns we see (Fig. 1). On the other hand, given that the eukaryotes diverged >1–2 billion years ago (Hedges 2002), it is likely that divergent evolution of RLK/*Pelle* homologs in eukaryotes has resulted in a situation where little signal is left for generating reliable phylogenies. If this is the case, why can we still find clear RLK/*Pelle* homologs in apicomplexans and metazoans? One intriguing possibility is that these apicomplexan RLK/*Pelles* may be products of horizontal transfer from a secondary symbiotic event involving green algae (Kohler et al. 1997), although this remains to be substantiated. In the case of metazoan RLK/*Pelles*, we speculate that conserved ancestral functions between animal and plant RLK/*Pelles* may have limited the degree of sequence divergence (see Sects. 3.1, 3.2, and 5.1).

In Viridiplantae, RLK/*Pelle* family members are found in multiple green algal species and all land plants (Fig. 1). Among chlorophyte algae, although no RLK/*Pelle* gene is found in *Ostreococcus tauri*, a green alga that is regarded as the smallest eukaryote (Courties et al. 1994), two RLCK genes are present in *Chlamydomonas reinhardtii* (Lehti-Shiu et al. 2009). In the charophyte lineage, which among green algae shares the closest common ancestor with land plants, multiple RLK/*Pelle* genes are found in *Closterium ehrenbergii*, a unicellular species, and *Nitella axillaris*, a multicellular charophyte (Sasaki et al. 2007). In land plants, the number of RLK/*Pelle* members rises sharply, with 329, 1,070, 610, and 1,192 genes found in moss, rice, *A. thaliana*, and poplar, respectively (Lehti-Shiu et al. 2009). This, together with the small number of RLK/*Pelle* genes in animal species, indicates that this gene family has experienced dramatic expansion specifically in the land plant lineage. In addition, given the presence of RLKs in charophytes, the receptor configuration likely arose before the divergence of charophytes from land plants. However, the charophyte RLKs were identified from transcript sequences. Thus, a more thorough phylogenetic study using the entire repertoire of RLK/*Pelle* genes in a charophyte genome will be necessary to assess whether there is a single or multiple origins of receptor configurations involving members of this family.

Why have RLK/*Pelle*s undergone such dramatic expansion only in plants? It is clear that different receptor kinase families have expanded in different lineages. In a comparative analysis of tyrosine kinases in animal species ranging from *Caenorhabditis elegans* to human, the tyrosine kinase family was found to have expanded substantially over the course of metazoan evolution (Shiu and Li 2004). In plants the expanded receptor kinase family is the RLK/*Pelle* family, and brown algae and oomycetes each contain phylogenetically distinct families of receptor kinases (Cock et al. 2010). Therefore, it is likely that kinases in different families were paired with ECDs in different lineages, and expansion of distinct receptor kinases were selected for independently due to the adaptive advantage conferred by the ability to perceive extracellular signals.

3 Evolution of RLK/*Pelle* Domain Content

3.1 RLK/Pelle *Domain Content Diversity and the Creation of Receptor Chimera*

The RLK/*Pelle* family can be divided into several subfamilies based on phylogenetic relationships between kinase domains (Shiu and Bleecker 2001a), and RLK/*Pelle* genes with related kinase domains almost always have the same type of ECD (Shiu and Bleecker 2001a, b). The diversity of plant RLK extracellular regions is similar to the domain complexity seen in animal receptor kinases (Cock et al. 2002), and this diversity has two implications. The first is the importance of RLKs in

perceiving a wide range of extracellular signals. Supporting this, multiple RLKs have been shown to directly bind to protein, lipid, polysaccharides, and other molecular ligands, both self and nonself (see Chapters "Receptor Ligands in Development" and "Phosphorylation and RLK signaling"). The second implication is that, because of the utility of transmembrane signaling, RLK/*Pelle* gene fusions were repeatedly selected for over the course of plant evolution. The fact that chimeric RLKs containing nonnative ECDs activate the same downstream signaling pathways as the native ECD (He et al. 2000; Albert et al. 2010; Brutus et al. 2010) illustrates how naturally occurring novel chimeric RLKs may have been created.

In plants, a large number of secreted or membrane spanning proteins have similar domain content as RLKs (Shiu and Bleecker 2003; Fritz-Laylin et al. 2005). These RLK-like proteins, collectively named Receptor-Like Proteins (RLPs), have a signal sequence and transmembrane domain but no intracellular kinase domain. Multiple RLPs have been shown to function with RLKs to regulate development and defense response. The RLKs *CLAVATA1* and *CORYNE* interact with the RLP *CLAVATA2* to regulate meristem development (Jeong et al. 1999; Muller et al. 2008). The RLP *TOO MANY MOUTHS* is proposed to form a complex with *ERECTA* family members to regulate stomatal patterning (Shpak et al. 2005). The rice LysM domain-containing RLP, *CEBiP*, and RLK, *OsCERK1*, transiently form a complex when treated with chitin oligosaccharides and activate defense response pathways (Shimizu et al. 2010). The similarities between RLPs and RLKs and their functional relationships are consistent with the possibility that RLKs with novel domain configurations may have been created through fusions between existing RLPs and RLKs/RLCKs. In addition, most RLPs, which are secreted or membrane spanning proteins, are likely integral components of extracellular signaling networks. Fusions between ancestral RLPs and RLK/*Pelle* kinases could therefore have led to novel signal transduction pathways by linking ligand perception to different downstream kinase targets. Alternatively, fusions may simply have occurred between RLPs and RLK/*Pelle*s that were already components of the same signaling networks.

The interaction between receptors lacking kinase domains and cytoplasmic kinases is known in animal systems; *D. melanogaster Pelle* is a cytoplasmic kinase that is part of signaling networks involving Toll, a transmembrane receptor without a kinase domain, mediating both innate immunity and development (Hecht and Anderson 1993; Shelton and Wasserman 1993; Belvin and Anderson 1996). Similarly mammalian IRAKs are parts of innate immunity signaling networks involving Toll-Like Receptors (Flannery and Bowie 2010). Multiple plant RLK/*Pelle* members are involved in innate immunity (Boller and Felix 2009). Thus, the innate immunity function of some members in the RLK/*Pelle* family is likely an ancestral trait. Given that the receptor configuration must arise from a fusion between an RLP and an RLCK, it is plausible that these RLKs with innate immunity functions were originally RLPs and RLCKs that fused together later on.

3.2 Rate of Domain Gain and Loss

Domain content diversity among RLK/*Pelle* family members clearly points to repeated creation of new RLKs (Shiu and Bleecker 2001a). Comparative genomic studies of RLK/*Pelle* families also provide support for multiple domain gain and loss events during flowering plant evolution. For example, after the poplar lineage separated from the *A. thaliana* lineage 100–120 million years ago (Tuskan et al. 2006), a novel poplar and a novel *A. thaliana* RLK were created by the fusion of a glycosyl hydrolase 18 (chitinase domain; Perrakis et al. 1994) domain and a glycerophosphodiester phosphodiesterase domain (GDPD; Santelli et al. 2004), respectively. In the rice lineage after its separation from dicots, two novel RLKs were generated involving a Pathogenesis-Related Protein-1/Sperm-Coating Glyco-protein domain (Szyperski et al. 1998) and a CHASE domain (Anantharaman and Aravind 2001; Mougel and Zhulin 2001).

There are also multiple examples where existing RLK ECDs were replaced by protein domains that are novel to the RLK/*Pelle* family (Shiu et al. 2004; Lehti-Shiu et al. 2009). In addition, several RLK ECDs are associated with more than one RLK subfamily. In these cases, based on the phylogenetic relationships between subfamilies, these common ECDs were likely acquired independently (Shiu et al. 2004; Lehti-Shiu et al. 2009). It is possible that the signaling mechanisms between subfamilies where domain swapping took place were similar and thus the swapped RLKs play the same roles. Alternatively, it is possible the swapping allowed an alternative signaling route for perceiving the same kinds of stimuli, assuming that different RLK subfamilies have distinct downstream signaling networks. As we learn more about RLK/*Pelle* functions, comparative analyses across RLK/*Pelle* subfamilies will allow us to distinguish between these two possibilities. In many cases, the domains gained have been implicated in biotic stress response, suggesting the selective pressure for proper response to biotic agents may have contributed to their retention (Shiu et al. 2004; Lehti-Shiu et al. 2009). Consistent with this view, a recent study has demonstrated that a gain of function mutation in the Arabidopsis LRK10L kinase containing the novel GDPD domain leads to constitutive activation of defense responses (Bi et al. 2010).

In addition to domain gain, domain loss has occurred during the evolution of the RLK/*Pelle* family. In *A. thaliana*, 28 members of RLK subfamilies have likely lost their ECDs based on phylogenetic analysis, leaving only an intracellular kinase domain typical of RLCKs. It is also likely that RLK/*Pelle* genes have lost kinase domains, leading to the generation of RLPs. This is consistent with the observation that some RLPs are located in close proximity to RLKs in their chromosomal positions (Shiu and Bleecker 2003). It was recently found that a *Drosophila* protein, Tube, which lacks a kinase domain, is evolutionarily related to mammalian *IRAK4* (Towb et al. 2009). Although it lacks kinase activity, *Tube* still functions in Toll signaling pathways as a scaffold protein. It is conceivable that, similarly, some RLPs may be derived from RLKs through kinase domain losses but still function in

the same signaling pathways as scaffold proteins similar to *Tube*. This can potentially be resolved through detailed phylogenomic analysis of RLP and RLK ECDs.

Although the diversity of RLKs is clearly the consequence of repeated innovation involving RLK/*Pelle* members, this does not mean that novel RLK configurations arose all that frequently. Among the RLK/*Pelle* subfamilies defined based on kinase phylogeny, 77% were established prior to the divergence of the vascular plant lineage from nonvascular plants ~450–700 million years ago (Hedges 2002; Lehti-Shiu et al. 2009). In addition, there are very few lineage-specific RLK/*Pelle* subfamilies. In a comparative analysis of the RLK/*Pelle* family in moss, rice, poplar, and *A. thaliana*, only two moss-specific and one poplar-specific subfamily were identified (Lehti-Shiu et al. 2009). Many RLK/*Pelle* subfamilies are also represented in liverwort (Sasaki et al. 2007). This suggests that most RLK/*Pelle* subfamilies were established very early in land plant evolution, well before the divergence of the land plant lineage ~450–700 million years ago. While domain fusions involving members of the RLK/*Pelle* family may have happened frequently early in land plant evolution, the rate of domain gain is very low. There have been only 12 instances of domain gain or swapping in three flowering plant lineages over the last 150 million years (Lehti-Shiu et al. 2009). The major reason for substantial differences in RLK/*Pelle* family size in plants is not due to the presence of many different RLKs with different domain content but due to differential expansion of existing subfamilies (Lehti-Shiu et al. 2009).

4   Expansion of the Plant RLK/*Pelle* Family

4.1   Dramatic Expansion in the Land Plant Lineage

As discussed in Sect. 2, there are zero or very few RLK/*Pelle* members in most eukaryotes except land plants, and in Viridiplantae, only two RLCKs are present in the chlorophyte alga *C. reinhardtii*. Thus, the small RLK/*Pelle* family sizes in animals and green algae are taken as evidence that the RLK/*Pelle* family likely was very small before the chlorophyte lineage diverged from land plants and related charophyte algae ~1 billion years ago (Hedges 2002). On the other hand, a charophyte alga, *C. ehrenbergii*, has at least 14 RLK/*Pelle*s, including an RLK (Sasaki et al. 2007). Note that charophyte algae belong to a polyphyletic group (Qiu 2008). *C. ehrenbergii* is in the family Zygnemophyceae that is not as close to land plants as those in the Charophyceae family such as *Chara* (Lewis and McCourt 2004). Thus, it will be of great interest to determine the abundance of RLK/*Pelle* genes in Charophyceae algae. The RLK/*Pelle* family expanded continuously over the course of land plant evolution; the early diverging bryophyte species, *Physcomitrella patens* (moss), has 329 RLK/*Pelle*s (Lehti-Shiu et al. 2009) and there are 610 RLK/*Pelle* genes in *A. thaliana* (Shiu and Bleecker 2001a), and over a thousand in rice (Shiu et al. 2004) and poplar (Lehti-Shiu et al. 2009). Given that

there were likely very few RLK/*Pelle* genes in the green algal-land plant common ancestor, the ancestral RLK/*Pelle* genes have been amplified greatly in the land plant lineage, particularly in vascular plants.

Based on a comparative analysis of RLK/*Pelle* protein sequences from moss and three flowering plants, it is estimated that the RLK/*Pelle* family has expanded by ~24% in the moss lineage since its divergence from the flowering plant lineage (Lehti-Shiu et al. 2009). In contrast, the expansion of the RLK/*Pelle* family in flowering plant lineages ranges from 93% to 212%. Based on a global comparison of plant gene families, the kinase superfamily, in particular, has undergone a much higher degree of expansion than nearly all gene families (Hanada et al. 2008). Given that the number of other kinases is ~400 in algae and land plants, the much higher degree of expansion seen in the kinase superfamily is mostly a consequence of expansion of the RLK/*Pelle* family (Lehti-Shiu et al. 2009). The expansion seen in the RLK/*Pelle* family is not only a result of many gene duplication events involving this gene family but also because many duplicates were retained. Because most duplicate genes become pseudogenes rather quickly (Li 1983), the much higher degree of expansion in the RLK/*Pelle* family relative to most other families in plants is mostly a consequence of differential retention.

Differential expansion among land plants is also apparent at the RLK/*Pelle* subfamily level (Lehti-Shiu et al. 2009). While some subfamilies have remained relatively constant in size, others have undergone pronounced differential expansion. For example, there are two C-LEC subfamily members in moss and only one in rice, poplar, and *A. thaliana*. In contrast, the rice WAK subfamily, which has over 114 members, is six times larger than in *A. thaliana*. Strikingly the rice and poplar SD1 subfamilies are more than ten times larger than in *A. thaliana*. Similar numbers of RLK/*Pelle* subfamily members in different species do not necessarily indicate that family sizes were established prior to divergence of these lineages, however. For example, the DUF26 subfamily expanded after the divergence of vascular plants, and there are similar numbers of DUF26 RLKs in poplar, *A. thaliana,* and rice. However, based on the phylogenetic relationships of DUF26 subfamily members in these plant species, the similar DUF26 subfamily sizes in these different lineages are mainly due to parallel, lineage-specific expansion and frequent gene loss (Shiu et al. 2004; Lehti-Shiu et al. 2009).

4.2   Duplication Mechanisms That Contribute to the RLK/Pelle Family Expansion

Expansion of a gene family involves a combination of gene duplication and subsequent retention. Thus, the dramatic expansion of the RLK/*Pelle* family in the land plant lineage and substantial variation in differential expansion among RLK/*Pelle* subfamilies across species prompt two important questions. The first is how these RLK/*Pelle*s were duplicated. The second question is whether these

RLK/*Pelle* duplicates were retained preferentially. The latter question will be discussed in Sect. 5. Compared to other eukaryotes, plants are unique in that whole genome duplications occurred much more frequently (Lockton and Gaut 2005; Cui et al. 2006). Together with other duplication mechanisms, such as tandem duplication, segmental duplication, and replicative transposition, a very large number of duplicate genes, in all gene families, have been generated in the long evolutionary history of land plants. Our current understanding is that the RLK/*Pelle* family has undergone dramatic expansion in land plant species and that this expansion is mainly a consequence of high rates of duplicate retention. However, substantial variation in retention rates within the family is revealed upon closer examination of subfamilies and orthologous groups. Interestingly, many RLK/*Pelle* genes are tandemly duplicated. In fact, 33% of the RLK/*Pelle* family in *A. thaliana* is found in tandem repeats (Lehti-Shiu et al. 2009). In rice and poplar, which have nearly double the number of RLK/*Pelle* genes, the proportion of members in tandem repeats is even higher (50% and 39%, respectively; Lehti-Shiu et al. 2009). In several studies, protein coding genes containing kinase domains have been shown to be overrepresented in tandem repeats (Rizzon et al. 2006; Tuskan et al. 2006; Hanada et al. 2008). Consistent with these studies, RLK/*Pelle* genes in orthologous groups that have expanded tend to be tandem genes (Lehti-Shiu et al. 2009), indicating that tandem duplication contributed more significantly to lineage-specific expansion of the RLK/*Pelle* family than other duplication mechanisms, such as whole genome duplication, combined.

Unequal crossovers, or tandem duplication, can change the content of tandem clusters dramatically (for example, resulting in multiple linked copies in one homologous chromosome and none in the other) and quickly (taking place every generation) (Leister 2004; Reams and Neidle 2004). This provides a mechanistic explanation for why RLKs are among the most variable gene families among *A. thaliana* accessions (Clark et al. 2007; Cao et al. 2011). Unequal crossover is most effective between genomic regions that are highly similar either because they are derived from recent duplication or have experienced gene conversion. Thus, the contribution of tandem duplication to RLK/*Pelle* family expansion is mostly on relatively young RLK/*Pelle* duplicates. We should also emphasize that other duplication mechanisms likely contributed to expansion as well, just that there is an overabundance of relatively young RLK/*Pelle* genes derived from tandem duplication. Based on the duplicated blocks that were derived from the most recent whole genome duplication in the *A. thaliana* lineage ~25–40 million years ago (Blanc et al. 2003), 63 RLK/*Pelle* gene pairs were retained (Lehti-Shiu et al. 2009). During the same time period, there were likely many more tandem duplicate RLK/*Pelle*s generated but lost. It is difficult to ascertain how the RLK/*Pelle* family was established early in land plant evolution because genomic features that allow us to discern duplication mechanisms quickly degenerated over time. In addition, it is not clear how mechanisms other than tandem and whole genome duplications have contributed to the RLK/*Pelle* family expansion. For example, retrogenes have contributed significantly to the gene content in animals (Pan and Zhang 2009) and in plants (Benovoy and Drouin 2006; Wang et al. 2006). Particularly

in repeat-rich plant genomes, it will be important to evaluate how retrotransposition contributed to RLK/*Pelle* family expansion.

Based on results of comparative genomic studies (Shiu et al. 2004; Hanada et al. 2008; Lehti-Shiu et al. 2009), it is clear that the RLK/*Pelle* family, together with the F-box gene family (Hua et al. 2011), has undergone the most dramatic lineage-specific expansion in land plants. Does this dramatic lineage specific expansion reflect the adaptive consequence of RLK/*Pelle* duplication? Or is it possible that the success of this gene family merely reflects a rare case of neutral evolution that can be explained by the genome drift hypothesis (Nozawa and Nei 2007)? To facilitate our discussion on why RLK/*Pelle* duplicates may be retained, we will first provide a short introduction of RLK/*Pelle* functions and how these functions are correlated with mechanisms of duplication and patterns of expansion in this gene family. For more details on RLK/*Pelle* functions, the readers are referred to in-depth reviews (Morillo and Tax 2006; Berger 2009; Tor et al. 2009; Zhao 2009; Li 2010; Postel et al. 2010; Rowe and Bergmann 2010) and other chapters in this volume.

5 Why Has the RLK/*Pelle* Family Expanded in the Land Plant Lineage

5.1 *Functions of RLK/Pelle Duplicates*

While the functions of most of the RLKs in plants remain unknown, much progress in elucidating RLK signaling networks has been made in recent years. Like the *Pelle* and IRAK genes, several plant RLK/*Pelle* family members have been shown to function in innate immunity as pattern recognition receptors (Albert et al. 2010). Several additional RLK/*Pelle*s function in defense response, binding pathogen components directly or acting in downstream pathways (Dodds and Rathjen 2010). A smaller group of RLKs have also been implicated in abiotic stress response (Yang et al. 2010; Sivaguru et al. 2003; Osakabe et al. 2005). In addition to plant stress response, by far the most RLKs with known functions are involved in some aspect of plant development regulating, for example, meristem size, organ identity, and cell-type specificity (reviewed in De Smet et al. 2009). Although most RLK/*Pelle*s with known functions are either involved in stress response or development regulation, recent studies indicate some RLK/*Pelle*s have more than one defined role. For example, BAK1 functions in both innate immunity and brassinosteroid signaling pathways, forming complexes with FLS2 and the BRI1 brassinosteroid receptor (reviewed in Chinchilla et al. 2009).

Members of a subfamily sometimes, but not always have related functions. One particularly well-studied subfamily is the LRR-XI family, of which *CLAVATA1* (*CLV1*) is a member. In *A. thaliana CLV1* functions in the meristem to restrict proliferation and promote differentiation (Brand et al. 2000). The closely related *BAM* RLKs are expressed in different regions of the meristem and promote stem cell

maintenance (DeYoung et al. 2006). The *CLV1* signaling pathway also appears to be conserved in monocots; mutations in the maize *CLV1* ((Taguchi-Shiobara et al. 2001; Bommert et al. 2005) and the rice *CLV1* (Suzaki et al. 2004) orthologs result in inflorescence meristem proliferation. However, not all LRR-XI family members function in meristems. In legumes, the orthologous LRR-XI subfamily members *NARK* from soybean, *HAR* and *Klavier* from *Lotus japonicus*, and *SUNN* from *Medicago truncatula* regulate nodule number but do not have meristem phenotypes when mutated (Krusell et al. 2002; Searle et al. 2003; Schnabel et al. 2005). In *A. thaliana,* the LRR RLK XI family members *PEPR1* and *PEPR2* recognize damage associated molecular patterns and activate innate immunity pathways (Yamaguchi et al. 2006; Krol et al. 2010; Yamaguchi et al. 2010). Clearly, different subfamily members have been recruited to play different roles during evolution. In addition, members from different subfamilies can have similar functions. For example, *ACR4*, a member of the *Crinkly4*-like subfamily, has a function analogous to *CLV1* in the root meristem, and despite having a different ECD, binds a similar ligand (Stahl et al. 2009).

5.2   Co-option of RLK/Pelle *Family Members for Plant–Microbe Interactions*

Co-option is the new use of existing genes, organs, or biological structures through natural selection, particularly, adaptive evolution (True and Carroll 2002). For duplicate genes, it has long been hypothesized that one copy retains the original role and the other copy may "neofunctionalize," i.e., take on novel functions (Ohno 1970). In this case, the duplicate copy can be retained if the novel function confers fitness advantage. Co-option of existing RLK functions has been shown to be important for the evolution of symbiosis. For example, given that nodulation is a legume-specific trait and that many LRR-XI members are involved in aspects of meristem development, the legume LRR-XI members discussed in the previous section are likely examples of co-option. Another example is the LysM subfamily whose members bind to chitin and are found in many plant species (Zhang et al. 2007; Buist et al. 2008). In legumes this family recognizes nod factors required for symbiosis, suggesting that nod recognition RLKs were recruited from existing chitin binding proteins (Shimizu et al. 2010). Nodulation recognition factors belonging to the LRR-I family were also co-opted to act in different nodulation pathways. SYMbiosis Receptor-like Kinase (*SYMRK*) and Nodulation Receptor Kinase (*NORK*) perceive lipo-chitooligosaccharide nodulation factors that are required for bacterial and fungal symbiosis (Endre et al. 2002; Stracke et al. 2002). *SYMRK* is required for symbiosis with mycorrhizal fungi as well as the formation of nitrogen fixing bacteria root nodule symbioses, which are restricted to legumes. It is thought that root nodule symbiosis likely evolved from the already existing arbuscular mycorrhizal symbiotic pathway (Soltis et al. 1995). Recently,

it was found that *SYMRK* played an important role in the evolution of root nodule symbiosis. While many species contain *SYMRK* orthologs and form arbuscular mychorriza, only legumes that have longer versions of the SYMRK gene are capable of nitrogen fixation (Markmann et al. 2008).

Although co-option and neofunctionalization may explain the retention of RLK/*Pelle* duplicates involved in symbiosis, there is little evidence that other RLK/*Pelle* duplicates were retained due to the acquisition of novel functions. In addition to neofunctionalization, the retention of duplicate genes can be attributed to subfunctionalization (Force et al. 1999; Lynch et al. 2001) where the ancestral functions of an RLK/*Pelle* were partitioned among the duplicate copies. If the partitioned "subfunctions" were all important, then both duplicate copies must have been retained even though the presence of both copies did not confer selective advantage over the ancestral gene. As more knowledge of RLK/*Pelle* functions accumulates, we will have a better understanding of the relative importance of neofunctionalization and/or subfunctionalization on RLK/*Pelle* duplicate retention.

5.3 Relationship Between RLK/Pelle *Function and Duplication Mechanism*

The RLK/*Pelle* members can be classified into multiple subfamilies, and these subfamilies show dramatically different degrees of expansion over the course of land plant evolution. Because many RLK/*Pelle* genes have been implicated in biotic stress, mostly on the basis of expression studies (Wrzaczek et al. 2010; Chae et al. 2009; Lehti-Shiu et al. 2009; Postel et al. 2010), one hypothesis explaining RLK/*Pelle* family expansion is that the selective pressure to properly perceive rapidly changing environmental signals has been one of the driving forces behind expansion. This notion is consistent with the finding that the RLK/*Pelle* family has a significantly overrepresented number of biotic stress responsive genes compared to all *A. thaliana* genes (Lehti-Shiu et al. 2009) and that many RLK/*Pelle*s are responsive to abiotic stress and hormone/chemical treatments (Chae et al. 2009). Furthermore, several RLK/*Pelle* subfamilies have significantly overrepresented numbers of genes that are responsive to particular stress conditions. For example, the DUF26, L-LEC, LRR-I, LRR-VIII-2, LRR-Xb, RLCK-VIIa, SD1, SD-2b, WAK, and WAK_LRK10L-1 subfamilies are enriched in genes that are responsive to a number of biotic stresses, and these families are also the ones that have undergone expansion (Lehti-Shiu et al. 2009).

There is evidence that lineage-specific expansion of stress-responsive RLKs has been driven mainly by tandem duplication. Among plant genes, duplicates generated by lineage-specific tandem duplication are more likely to function in stress response (Hanada et al. 2008). This is also true for tandem RLK/*Pelle* genes, which are significantly more likely to be responsive to biotic stress than nontandem genes (Lehti-Shiu et al. 2009). One explanation for the relationship between tandem

duplication and stress responsiveness is that tandem duplications allow rapid changes in gene content over just a few generations. This can generate a higher degree of diversity than other duplication mechanisms such as whole genome duplication. In fact, the NBS-LRR family of resistance genes which recognize pathogen effectors and trigger disease resistance pathways have a similar pattern of expansion as RLK/*Pelle* genes, with varying numbers of genes found between species and even between accessions of the same species (Zhang et al. 2010).

The expansion of stress-responsive RLKs by tandem duplication is supported when looking at RLKs with known functions. RLKs with roles in development tend not to be in tandem clusters whereas those with roles in defense tend to be in tandem repeats (Shiu et al. 2004). Interesting exceptions to this are the LRR-XII FLS2 and EFR pattern recognition receptors in *A. thaliana*. Members of the LRR-XII subfamily have broadly conserved functions, with subfamily members in different species activating similar transduction pathways. FLS2 can be found in tomato and rice (Robatzek et al. 2007; Takai et al. 2008), and the signaling pathways induced by binding to LRR-XII receptors are conserved. Interestingly, while there are only ten LRR-XII subfamily members in *A. thaliana*, there are ten times as many members in rice and poplar. It was reasoned that the relatively small number of members in *A. thaliana* may be due to the fact that since PAMP receptors recognize conserved features of microbes, diversification of this subfamily is not selected for (Tang et al. 2010). Another explanation is that the differential expansion of LRR-XII in poplar and rice is the consequence of intense selective pressure, perhaps due to their involvement in a recent or ongoing arms race with yet to be identified biotic agents. Note that the previous two explanations assume natural selection in the form of adaptive evolution is the driving force behind RLK/*Pelle* retention. Another possibility, however, is that differences in numbers between species could be due to random "genomic drift" and that expansion does not necessarily imply an adaptive advantage (Nozawa and Nei 2007; Nei et al. 2008).

5.4 Is RLK/Pelle *Family Expansion the Result of Neutral or Adaptive Evolution?*

Based on our understanding of the differences and similarities between the RLK/*Pelle* families within and between land plant species, it is now possible to speculate on the contribution of neutral and adaptive evolution to the expansion of this gene family. Gene duplication is one type of mutation that is thought to occur randomly. Thus, in general, larger gene families tend to have more lineage-specific duplicates compared to smaller gene families in plants (Hanada et al. 2008; Zou et al. 2009). Nonetheless, duplicated genes are not necessarily retained. Thus, the expansion of a gene family is controlled by how often its members are duplicated and retained. The central question is whether retention is due to natural selection or some neutral processes that occur randomly. The latter possibility is the central tenet of the

genome drift hypothesis, which postulates that the expansion and shrinkage of gene families can be explained by random duplication and inactivation events analogous to how genetic drift affects allelic frequencies (Nozawa and Nei 2007; Nei et al. 2008; Nei et al. 2010). The metazoan olfactory receptor (OR) family has been used as an important example of genome drift (Nozawa and Nei 2007). The OR family has three important features consistent with a family that may have experienced genome drift: (1) extensive lineage-specific differences between metazoan species, (2) highly variable within species, (3) a large number of OR pseudogenes littering the human genome. Among plant gene families, the F-box family is the most similar to the metazoan OR family in these three aspects (Hua et al. 2011). In the case of RLK/*Pelle*, there are some similarities but the pattern is rather complicated.

First of all, as detailed in Sect. 4.1, there is extensive lineage-specific expansion in the RLK/*Pelle* family among flowering plants. RLK/*Pelle* members also rank highly in the degree of intraspecific variation (Clark et al. 2007; Cao et al. 2011). However, although there are a large number of RLK/*Pelle* pseudogenes in *A. thaliana* and rice, the RLK/*Pelle* family in fact has significantly fewer pseudogenes compared to most other gene families (Zou et al. 2009). In contrast, the LRR-R genes and F-box genes both have significantly more pseudogenes than most plant gene families. Thus, neutral processes such as genome drift do not entirely account for the degree of RLK/*Pelle* family expansion. At least some of the duplicated copies were retained due to natural selection. Consistent with this speculation, the co-opted RLK/*Pelle*s involved in symbiotic associations are clearly examples of selection, in these cases through adaptive evolution. And results from comparative genomics provide ample examples of ancient RLK/*Pelle* duplicates that remain conserved across land plant species, arguing against retention solely due to neutral explanations (Shiu et al. 2004; Lehti-Shiu et al. 2009).

We should point out that some RLK/*Pelle* subfamilies evolve much like the OR family with extensive lineage-specific differences, substantial intraspecific variation, and a high pseudogene-to-functional gene ratio (Lehti-Shiu et al. 2009; Zou et al. 2009). Are these RLK/*Pelle*s still around due to genome drift? A significant number of these types of OR-like RLK/*Pelle*s are derived from tandem duplication and evolve quickly. In addition, multiple protein domains found in these RLK/*Pelle*s are involved in perceiving fungal or bacterial molecular patterns (see Sect. 3.2). Furthermore, an overrepresented number of RLK/*Pelle*s are differentially upregulated under biotic stress conditions. Taking all this information into consideration, the evolution of the RLK/*Pelle* family may be best explained by a mixture of drift and selection. It is conceivable that many RLK/*Pelle* duplicates are generated by chance. Over the course of land plant evolution, although many duplicates did not survive, some were selected for because of their ability to regulate novel developmental processes or mediate responses to symbiotic or pathogenic microbes. Many RLK/*Pelle*s with known functions are involved in developmental regulation and are clearly indispensible and conserved across species. In some cases the molecular patterns recognized by the selected RLK/*Pelle*s are not easily mutable, such as flagellin. In other cases, members of this gene family may engage in arms races with biotic agents much like the LRR-R genes. To

determine if the above scenario is true or not, we need more information on how RLK/*Pelle* ortholog functions have evolved in multiple land plant species. We also need a better handle on the roles of rapidly evolving RLK/*Pelle*s, particularly those in tandem repeats.

6   Conclusions

The timing and extent of RLK/*Pelle* family expansion suggests that this family has played a significant role in the evolution of land plants. Closer examination of the patterns of RLK/*Pelle* subfamily expansion has revealed that much expansion is lineage specific and driven by both whole genome and tandem duplication. Based on rapidly accumulating information on the functions of RLK/*Pelle* genes, a clearer picture of the selection pressures driving expansion is emerging. Most duplicate genes become pseudogenes within a few million years. However, because larger families tend to generate more duplicates, and the RLK/*Pelle* family is one of the largest in plants, some RLK/*Pelle* duplicates may still be around simply because many of them are generated and there is not sufficient time for their pseudogenization. Thus, some of the observed expansion is likely a consequence of completely neutral processes that do not involve natural selection.

Nonetheless, many RLK/*Pelle* genes are known to have roles in biotic stress signaling and have likely been co-opted for symbiotic interactions. This suggests that the pressure to perceive and respond to rapidly evolving biotic factors is a likely driving force behind expansion. Furthermore, the fact that many RLK/*Pelle* duplicates have survived over tens to hundreds of millions of years indicates that natural selection plays an important role in RLK/*Pelle* expansion. One of the biggest challenges to understanding the mechanisms that drive the expansion of this family is in identifying the roles of fast evolving, tandem RLK/*Pelle*s. In addition, although related RLK/*Pelle* sequences can function in similar pathways in different species, it is clear that RLK/*Pelle*s have been co-opted for new signaling roles. This makes extrapolations of functions based on studies from one species complicated. As more knowledge about RLK/*Pelle* functions becomes available, it can be combined with information about expansion, duplication mechanism, and conservation of family members across species to better understand the functional evolution of this large gene family.

References

Albert M, Jehle AK, Mueller K, Eisele C, Lipschis M, Felix G (2010) *Arabidopsis thaliana* pattern recognition receptors for bacterial elongation factor Tu and flagellin can be combined to form functional chimeric receptors. J Biol Chem 285:19035–19042

Anantharaman V, Aravind L (2001) The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors. Trends Biochem Sci 26:579–582

Belvin MP, Anderson KV (1996) A conserved signaling pathway: the Drosophila Toll-Dorsal pathway. Annu Rev Cell Dev Biol 12:393–416

Benovoy D, Drouin G (2006) Processed pseudogenes, processed genes, and spontaneous mutations in the Arabidopsis genome. J Mol Evol 62:511–522

Berger F (2009) Reproductive biology: receptor-like kinases orchestrate love songs in plants. Curr Biol 19:R647–R649

Bi D, Cheng YT, Li X, Zhang Y (2010) Activation of plant immune responses by a gain-of-function mutation in an atypical receptor-like kinase. Plant Physiol 153:1771–1779

Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Res 13:137–144

Boller T, Felix G (2009) A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu Rev Plant Biol 60:379–406

Bommert P, Lunde C, Nardmann J, Vollbrecht E, Running M, Jackson D, Hake S, Werr W (2005) Thick tassel dwarf1 encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase. Development 132:1235–1245

Brand U, Fletcher JC, Hobe M, Meyerowitz EM, Simon R (2000) Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. Science 289:617–619

Braun DM, Walker JC (1996) Plant transmembrane receptors: new pieces in the signaling puzzle. Trends Biochem Sci 21:70–73

Brutus A, Sicilia F, Macone A, Cervone F, De Lorenzo G (2010) A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. Proc Natl Acad Sci USA 107:9452–9457

Buist G, Steen A, Kok J, Kuipers OP (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. Mol Microbiol 68:838–847

Cao Z, Henzel WJ, Gao X (1996) IRAK: a kinase associated with the interleukin-1 receptor. Science 271:1128–1131

Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Müller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nature Genet 43:956–963

Chae L, Sudat S, Dudoit S, Zhu T, Luan S (2009) Diverse transcriptional programs associated with environmental stress and hormones in the Arabidopsis receptor-like kinase gene family. Mol Plant 2:84–107

Chinchilla D, Shan L, He P, de Vries S, Kemmerling B (2009) One for all: the receptor-associated kinase BAK1. Trends Plant Sci 14:535–541

Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds DA, Chen H, Frazer KA, Huson DH, Scholkopf B, Nordborg M, Ratsch G, Ecker JR, Weigel D (2007) Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science 317:338–342

Cock JM, Vanoosthuyse V, Gaude T (2002) Receptor kinase signalling in plants and animals: distinct molecular systems with mechanistic similarities. Curr Opin Cell Biol 14:230–236

Cock JM, Sterck L, Rouze P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury JM, Badger JH, Beszteri B, Billiau K, Bonnet E, Bothwell JH, Bowler C, Boyen C, Brownlee C, Carrano CJ, Charrier B, Cho GY, Coelho SM, Collen J, Corre E, Da Silva C, Delage L, Delaroque N, Dittami SM, Doulbeau S, Elias M, Farnham G, Gachon CM, Gschloess B, Heesch S, Jabbari K, Jubin C, Kawai H, Kimura K, Kloareg B, Kupper FC, Lang D, Le Bail A, Leblanc C, Lerouge P, Lohr M, Lopez PJ, Martens C, Maumus F, Michel G, Miranda-Saavedra D, Morales J, Moreau H, Motomura T, Nagasato C, Napoli CA, Nelson DR, Nyvall-Collen P, Peters AF, Pommier C, Potin P, Poulain J, Quesneville H, Read B, Rensing SA, RitterA RS, Samanta M, Samson G, Schroeder DC, Segurens B, Strittmatter M, Tonon T, Tregear JW, Valentin K, von Dassow P, Yamagishi T, Van de Peer Y, Wincker P (2010) The Ectocarpus genome and the independent evolution of multicellularity in brown algae. Nature 465:617–621

Courties C, Vaquer A, Troussellier M, Lautier J, Chrétiennot-Dinet MJ, Neveux J, Machado C, Claustre H (1994) Smallest eukaryotic organism. Nature 370:255

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, Albert VA, Ma H, dePamphilis CW (2006) Widespread genome duplications throughout the history of flowering plants. Genome Res 16:738–749

De Smet I, Voss U, Jurgens G, Beeckman T (2009) Receptor-like kinases shape the plant. Nat Cell Biol 11:1166–1173

DeYoung BJ, Bickle KL, Schrage KJ, Muskett P, Patel K, Clark SE (2006) The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in Arabidopsis. Plant J 45:1–16

Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet 11:539–548

Endre G, Kereszt A, Kevei Z, Mihacea S, Kalo P, Kiss GB (2002) A receptor kinase gene regulating symbiotic nodule development. Nature 417:962–966

Flannery S, Bowie AG (2010) The interleukin-1 receptor-associated kinases: critical regulators of innate immune signalling. Biochem Pharmacol 80:1981–1991

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerate mutations. Genetics 151:1531–1545

Fritz-Laylin LK, Krishnamurthy N, Tor M, Sjolander KV, Jones JD (2005) Phylogenomic analysis of the receptor-like proteins of rice and Arabidopsis. Plant Physiol 138:611–623

Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu SH (2008) Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiol 148:993–1003

Handford PA, Baron M, Mayhew M, Willis A, Beesley T, Brownlee GG, Campbell ID (1990) The first EGF-like domain from human factor IX contains a high-affinity calcium binding site. EMBO J 9:475–480

He Z, Wang ZY, Li J, Zhu Q, Lamb C, Ronald P, Chory J (2000) Perception of barssinosteroids by the extracellular domain of the receptor kinase BRI1. Science 288:2360–2363

Hecht PM, Anderson KV (1993) Genetic characterization of tube and pelle, genes required for signaling between Toll and dorsal in the specification of the dorsal-ventral pattern of the Drosophila embryo. Genetics 135:405–417

Hedges SB (2002) The origin and evolution of model organisms. Nat Rev Genet 3:838–849

Hester G, Kaku H, Goldstein IJ, Wright CS (1995) Structure of mannose-specific snowdrop (Galanthus nivalis) lectin is representative of a new plant lectin family. Nat Struct Biol 2:472–479

Hinata K, Watanabe M, Yamakawa S, Satta Y, Isogai A (1995) Evolutionary aspects of the S-related genes of the Brassica self-incompatibility system: synonymous and nonsynonymous base substitutions. Genetics 140:1099–1104

Hua Z, Zou C, Shiu SH, Vierstra RD (2011) Phylogenetic comparison of F-Box (FBX) gene superfamily within the plant kingdom reveals divergent evolutionary histories indicative of genomic drift. PLoS One 6:e16219

Hunter T, Lindberg RA, Middlemas DS, Tracy S, van der Geer P (1992) Receptor protein tyrosine kinases and phosphatases. Cold Spring Harb Symp Quant Biol 57:25–41

Janssens S, Beyaert R (2003) Functional diversity and regulation of different interleukin-1 receptor-associated kinase (IRAK) family members. Mol Cell 11:293–302

Jeong S, Trotochaud AE, Clark SE (1999) The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinsae. Plant Cell 11:1925–1934

King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li W, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, Sequencing JG, Bork P, Lim WA, Manning G, Miller WT, McGinnis W, Shapiro H, Tjian

R, Grigoriev IV, Rokhsar D (2008) The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature 451:783–788

Kobe B, Deisenhofer J (1994) The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci 19:415–421

Kohler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJ, Palmer JD, Roos DS (1997) A plastid of probable green algal origin in Apicomplexan parasites. Science 275:1485–1489

Kohorn BD (2001) WAKs; cell wall associated kinases. Curr Opin Cell Biol 13:529–533

Krol E, Mentzel T, Chinchilla D, Boller T, Felix G, Kemmerling B, Postel S, Arents M, Jeworutzki E, Al-Rasheid KA, Becker D, Hedrich R (2010) Perception of the Arabidopsis danger signal peptide 1 involves the pattern recognition receptor AtPEPR1 and its close homologue AtPEPR2. J Biol Chem 285:13471–13479

Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczyglowski K, Duc G, Kaneko T, Tabata S, de Bruijn F, Pajuelo E, Sandal N, Stougaard J (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature 420:422–426

Lease KA, Walker JC (2006) The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. Plant Physiol 142:831–838

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

Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. Trends Genet 20:116–122

Lewis LA, McCourt RM (2004) Green algae and the origin of land plants. Am J Bot 91:1535–1556

Li WH (1983) Evolution of duplicate genes and pseudogenes. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer Associates, Sunderland, MA, pp 14–37

Li J (2010) Multi-tasking of somatic embryogenesis receptor-like protein kinases. Curr Opin Plant Biol 13:509–514

Lockton S, Gaut BS (2005) Plant conserved non-coding sequences and paralogue evolution. Trends Genet 21:60–65

Loris R, Hamelryck T, Bouckaert J, Wyns L (1998) Legume lectin structure. Biochim Biophys Acta 1383:9–36

Lynch M, O'Hely M, Walsh B, Force A (2001) The probability of preservation of a newly arisen gene duplicate. Genetics 159:1789–1804

Markmann K, Giczey G, Parniske M (2008) Functional adaptation of a plant receptor-kinase paved the way for the evolution of intracellular root symbioses with bacteria. PLoS Biol 6:e68

Miyakawa T, Sawano Y, Miyazono K, Hatano K, Tanokura M (2007) Crystallization and preliminary X-ray analysis of ginkbilobin-2 from *Ginkgo biloba* seeds: a novel antifungal protein with homology to the extracellular domain of plant cysteine-rich receptor-like kinases. Acta Crystallogr Sect F Struct Biol Cryst Commun 63:737–739

Morillo SA, Tax FE (2006) Functional analysis of receptor-like kinases in monocots and dicots. Curr Opin Plant Biol 9:460–469

Mougel C, Zhulin IB (2001) CHASE: an extracellular sensing domain common to transmembrane receptors from prokaryotes, lower eukaryotes and plants. Trends Biochem Sci 26:582–584

Muller R, Bleckmann A, Simon R (2008) The receptor kinase CORYNE of Arabidopsis transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. Plant Cell 20:934–946

Nei M, Niimura Y, Nozawa M (2008) The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. Nat Rev Genet 9:951–963

Nei M, Suzuki Y, Nozawa M (2010) The neutral theory of molecular evolution in the genomic era. Annu Rev Genomics Hum Genet 11:265–289

Nozawa M, Nei M (2007) Evolutionary dynamics of olfactory receptor genes in Drosophila species. Proc Natl Acad Sci USA 104:7122–7127

Ohno S (1970) Evolution by gene duplication. Springer, New York

Osakabe Y, Maruyama K, Seki M, Satou M, Shinozaki K, Yamaguchi-Shinozaki K (2005) Leucine-rich repeat receptor-like kinase1 is a key membrane-bound regulator of abscisic acid early signaling in Arabidopsis. Plant Cell 17:1105–1119

Pan D, Zhang L (2009) Burst of young retrogenes and independent retrogene formation in mammals. PLoS One 4:e5040

Pan CH, Lee EA, Chae YA, Kim SI (1999) Purification of chitinolytic protein from Rehmannia glutinosa showing N-terminal amino acid sequence similarity to thaumatin-like proteins. Biosci Biotechnol Biochem 63:1138–1140

Perrakis A, Tews I, Dauter Z, Oppenheim AB, Chet I, Wilson KS, Vorgias CE (1994) Crystal structure of a bacterial chitinase at 2.3 A resolution. Structure 2:1169–1180

Postel S, Kufner I, Beuter C, Mazzotta S, Schwedt A, Borlotti A, Halter T, Kemmerling B, Nurnberger T (2010) The multifunctional leucine-rich repeat receptor kinase BAK1 is implicated in Arabidopsis development and immunity. Eur J Cell Biol 89:169–174

Qiu Y-L (2008) Phylogeny and evolution of charophytic algae and land plants. J Syst Evol 46:287–306

Reams AB, Neidle EL (2004) Selection for gene clustering by tandem duplication. Annu Rev Microbiol 58:119–142

Rizzon C, Ponger L, Gaut BS (2006) Striking similarities in the genomic distribution of tandemly arrayed genes in Arabidopsis and rice. PLoS Comput Biol 2:e115

Robatzek S, Bittel P, Chinchilla D, Kochner P, Felix G, Shiu SH, Boller T (2007) Molecular identification and characterization of the tomato flagellin receptor LeFLS2, an orthologue of Arabidopsis FLS2 exhibiting characteristically different perception specificities. Plant Mol Biol 64:539–547

Rowe MH, Bergmann DC (2010) Complex signals for simple cells: the expanding ranks of signals and receptors guiding stomatal development. Curr Opin Plant Biol 13:548–555

Santelli E, Schwarzenbache R, McMullan D, Biorac T, Brinen LS, Canaves JM, Cambell J, Dai X, Deacon AM, Elsliger MA, Eshagi S, Floyd R, Godzik A, Grittini C, Grzechnik SK, Jaroszewski L, Karlak C, Klock HE, Koesema E, Kovarik JS, Kreusch A, Kuhn P, Lesley SA, McPhillips TM, Miller MD, Morse A, Moy K, Ouyang J, Page R, Quijano K, Rezezadeh F, Robb A, Sims E, Spraggon G, Stevens RC, van den Bedem H, Velasquez J, Vincent J, von Delft F, Wang X, West B, Wolf G, Xu Q, Hodgson KO, Wooley J, Wilson IA (2004) Crystal structure of a glycerophosphodiester phosphodiesterase (GDPD) from *Thermotoga maritima* (TM1621) at 1.60 A resolution. Proteins 56:167–170

Sasaki G, Katoh K, Hirose N, Suga H, Kuma K, Miyata T, Su ZH (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 401:135–144

Schnabel E, Journet EP, de Carvalho-Niebel F, Duc G, Frugoli J (2005) The *Medicago truncatula* SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. Plant Mol Biol 58:809–822

Searle IR, Men AE, Laniya TS, Buzas DM, Iturbe-Ormaetxe I, Carroll BJ, Gresshoff PM (2003) Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. Science 299:109–112

Sharon N, Lis H (2001) The structural basis for carbohydrate recognition by lectins. Adv Exp Med Biol 491:1–16

Shelton CA, Wasserman SA (1993) Pelle encodes a protein kinase required to establish dorsoventral polarity in the Drosophila embryo. Cell 72:515–525

Shimizu T, Nakano T, Takamizawa D, Desaki Y, Ishii-Minami N, Nishizawa Y, Minami E, Okada K, Yamane H, Kaku H, Shibuya N (2010) Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. Plant J 64:204–214

Shiu SH, Bleecker AB (2001a) Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proc Natl Acad Sci USA 98:10763–10768

Shiu SH, Bleecker AB (2001b) Plant receptor-like kinase gene family: diversity, function, and signaling. Sci STKE 113:re22

Shiu SH, Bleecker AB (2003) Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Plant Physiol 132:530–543

Shiu SH, Li WH (2004) Origins, lineage-specific expansions, and multiple losses of tyrosine kinases in eukaryotes. Mol Biol Evol 21:828–840

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 16:1220–1234

Shpak ED, McAbee JM, Pillitteri LJ, Torii KU (2005) Stomatal patterning and differentiation by synergistic interactions of receptor kinases. Science 309:290–293

Sivaguru M, Ezaki B, He ZH, Tong H, Osawa H, Baluska F, Volkmann D, Matsumoto H (2003) Aluminum-induced gene expression and protein localization of a cell wall-associated receptor kinase in Arabidopsis. Plant Physiol 132:2256–2266

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc Natl Acad Sci USA 92:2647–2651

Stahl Y, Wink RH, Ingram GC, Simon R (2009) A signaling module controlling the stem cell niche in Arabidopsis root meristems. Curr Biol 19:909–914

Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Szczyglowski K, Parniske M (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature 417:959–962

Suzaki T, Sato M, Ashikari M, Miyoshi M, Nagato Y, Hirano HY (2004) The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1. Development 131:5649–5657

Szyperski T, Fernandez C, Mumenthaler C, Wuthrich K (1998) Structure comparison of human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein P14a indicates a functional link between the human immune system and a plant defense system. Proc Natl Acad Sci USA 95:2262–2266

Taguchi-Shiobara F, Yuan Z, Hake S, Jackson D (2001) The fasciated ear2 gene encodes a leucine-rich repeat receptor-like protein that regulates shoot meristem proliferation in maize. Genes Dev 15:2755–2766

Takai R, Isogai A, Takayama S, Che FS (2008) Analysis of flagellin perception mediated by flg22 receptor OsFLS2 in rice. Mol Plant Microbe Interact 21:1635–1642

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 Sci 179:399–406

Tor M, Lotze MT, Holton N (2009) Receptor-mediated signalling in plants: molecular patterns and programmes. J Exp Bot 60:3645–3654

Tordai H, Banyai 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 46:163–167

Torii KU (2000) Receptor kinase activation and signal transduction in plants: an emerging picture. Curr Opin Plant Biol 3:361–367

Towb P, Sun H, Wasserman SA (2009) Tube Is an IRAK-4 homolog in a Toll pathway adapted for development and immunity. J Innate Immun 1:309–321

True JR, Carroll SB (2002) Gene co-option in physiological and morphological evolution. Annu Rev Cell Dev Biol 18:53–80

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapma J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Elli B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S,