Malectin/Malectin-like domain-containing proteins: A repertoire of cell surface molecules with broad functional potential

He Yang a, b, Dong Wang a, b, c, Li Guo b, d, Huairong Pan b, e, Robert Yvon a, b, Scott Garman a, b, Hen-Ming Wu a, b, c, *, Alice Y. Cheung a, b, c, *

a Department of Biochemistry and Molecular Biology, University of Massachusetts, USA
b Molecular and Cellular Biology Program, University of Massachusetts, USA
c Plant Biology Graduate Program, University of Massachusetts, Amherst, MA 01003, USA
d Faculty of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an 710049, China
e College of Biology, Hunan University, Changsha 410082, China

ARTICLE INFO

Keywords:
Malectin
Malectin-like domains (MLDs)
Carbohydrate binding modules
Duplication
Expansion of M/MLD-RLKs
Model and crop plants

ABSTRACT

Cell walls are at the front line of interactions between walled-organisms and their environment. They support cell expansion, ensure cell integrity and, for multicellular organisms such as plants, they provide cell adherence, support cell shape morphogenesis and mediate cell-cell communication. Wall-sensing, detecting perturbations in the wall and signaling the cell to respond accordingly, is crucial for growth and survival. In recent years, plant signaling research has suggested that a large family of receptor-like kinases (RLKs) could function as wall sensors partly because their extracellular domains show homology with malectin, a diglucose binding protein from the endoplasmic reticulum of animal cells. Studies of several malectin/malectin-like (M/ML) domain-containing RLKs (M/MLD-RLKs) from the model plant Arabidopsis thaliana have revealed an impressive array of biological roles, controlling growth, reproduction and stress responses, processes that in various ways rely on or affect the cell wall. Malectin homologous sequences are widespread across biological kingdoms, but plants have uniquely evolved a highly expanded family of proteins with Ml domains embedded within various protein contexts. Here, we present an overview on proteins with malectin homologous sequences in different kingdoms, discuss the chromosomal organization of Arabidopsis M/MLD-RLKs and the phylogenetic relationship between these proteins from several model and crop species. We also discuss briefly the molecular networks that enable the diverse biological roles served by M/ML-RLKs studied thus far.

Introduction

Malectins are conserved disaccharide-binding proteins of about 190 amino acid residues found in the lumen of the endoplasmic reticulum (ER) of animal cells (Schallus et al., 2008, Schallus et al., 2010). They associate with the oligosaccharyltransferase complex and bind to the diglucosylated intermediate of N-linked glycans early in the secretory pathway and target misfolded glycoproteins to degradation, thus playing an important role in protein quality control (Müller et al., 2010; Galli et al., 2011; Takeda et al., 2014; Tannous et al., 2014). In vitro, malectins bind diglucose, such as maltose (glucose-α-1,4 glucose) and nigerose (glucose-α-1,3 glucose), the latter its potential native ligand in the ER (Schallus et al., 2008). Plants have evolved exceptionally large families of receptor-like kinases (RLKs) and they have diversified into several distinct clades (Shiu and Bleecker, 2003; Shiu et al., 2004). The discovery of malectin revealed a group of plant RLKs as having malectin homologous regions in their extracellular domains (ECDs). As single-span transmembrane proteins, the ECDs of RLKs decorate the cell membrane-cell wall interface. The exterior of the cell is often considered topologically equivalent to the ER lumen in that membrane proteins facing the ER lumen are exposed to extracellular environment (Cooper and Hausman, 2004), which, in walled cells, are enriched in complex carbohydrates. These properties and the functions of the first two studied Malectin/Malectin-like (M/ML) domain RLKs (M/MLD-RLKs), THESEUS1 (THE1) and FERONIA (FER), from the model plant Arabidopsis thaliana (Hematy et al., 2007; Escobar-Restrepo et al., 2007),...
provoked widespread speculations that the ML sequences might interact with cell wall polysaccharides and that these proteins might serve as wall sensors (Hematy and Hofte, 2008; Cheung and Wu, 2011; Boisson-Dernier et al., 2011). Studies since then have demonstrated that some M/MLD-RLKs are crucial for plant growth and survival and regulate diverse processes, including the capacity to mediate cell wall-related phenomena. Given the intense interest on M/MLD-RLKs in recent years, we undertake a discussion here regarding the distribution of malectin homologous sequences from bacteria to higher plants and explore how knowledge from carbohydrate binding modules (CBMs) in bacterial proteins and M/ML domain-containing proteins in other species might inform about the evolution of plant M/MLD-RLKs. We also examine the chromosomal organization of a subset of M/MLD-RLKs from Arabidopsis and the phylogeny of related protein families from several model and crop plant species. The functions of some of these M/MLD-RLKs and the molecular interactions that they engage in are discussed briefly to provide some biological contexts. We refer readers to recent reviews for deeper discussions (Galindo-Trigo et al., 2016; Frank et al., 2018; Doblas et al., 2018)

Widespread presence of malectin homologous sequences and their expansion in plants

The malectin globular core is structurally most similar to several non-catalytic CBMs present in bacterium-secreted glycolytic enzymes found in the plant cell wall (Schallus et al., 2008; Gilbert et al., 2013; Duan et al., 2016). CBMs interact with non-substrate carbohydrate moieties from glycan or short oligosaccharide chains in the cell wall. This might augment the proximity between the catalytic domain of the enzymes and their substrate carbohydrates and potentiate their glycolytic activity (Crouch et al., 2016). In animal cells the binding of malectin to diglycosylated glycans early in the secretory pathway might recruit other proteins for glycol-modification or protein quality control (Tannous et al., 2014). It is possible that the carbohydrate-binding propensity in malectins and ML sequences, whether acting alone or as part of a larger protein, could facilitate processes occurring in a context involving carbohydrates.

Malectin homologous sequences are widespread throughout biological kingdoms (Fig. 1A). The Pfam database for protein families (http://pfam.xfam.org) (Mistry et al., 2021) identifies numerous malectin homologous sequences and classifies them into malectin (PF11721) and ML (PF12819) subfamilies. Both families are grouped under the large Galactose-binding domain-like superfamily (CL0202) containing ~70 families of proteins with beta sandwich domains and a jelly roll topology [https://pfam.xfam.org/clan/CL0202]. To avoid confusion with animal malectins, we shall use “malectin domains” to designate the PF11721 domains from hereon. Relative to malectin domains, MLs are generally longer and more variable in lengths, some have...
two complete malectin homologous sequences (Fig. 1B). More than 5000 malectin domains have been identified (Fig. 1A) with almost 3400 from *Viridiplantae* (land plants), distributing across 121 species analyzed whereas on the average only one to eleven malectin domains are found in each of the other species examined. Fungi, evolutionarily close to animals, however lack malectin homologous sequences. The ML family has emerged uniquely in plants, became highly expanded, averaging ~60 per species (Fig. 1A), and diversified into several distinct clades (Fig. 2A). The expansion of malectin homologous sequences in plants is remarkable and, like the plant RLK superfamily (Shiu et al., 2004), could have evolved to generate expanded capacity to meet the functional needs of land plants, sessile but in an always-changing environment.

The malectin domains found in bacteria, animals and plants have distinct characteristics (Fig. 1B). An N-terminal signal peptide targets animal malectins to the secretory pathway, and a short C-terminal transmembrane helix tethers its carbohydrate-binding core to the ER membrane (Schallus et al., 2008). In bacteria, malectin domains are usually within glycosyl hydrolases. In prokaryotic Archaea and other eukaryotic taxonomic groups such as the *Rhodophyta*, *Alveolata* and *Choofoflagellate*, malectin domains are embedded in different protein contexts. Most plant malectin homologous sequences are in the ECDs of protein kinases. These differences suggest that organisms in these taxa had experienced lineage-specific divergence in malectin homologous sequence-containing proteins. Association of malectin homologous sequences with diverse protein contexts could be a result of different functional adaptations. Animal malectin functions in a relatively simple carbohydrate environment and interact primarily with intermediates of glycoproteins undergoing modification and quality control checkpoints in the ER (Schallus et al., 2008; Tannous et al., 2014). The malectin domains in bacterial proteins might function as CBMs of cell wall hydrolases to potentiate their catalytic activities (Gilbert et al., 2013; Duan et al., 2016; Crouch et al., 2016). When linked with domains that mediate protein–protein interactions, such as Kelch motifs (Adams et al., 2000) (Fig. 1B) and leucine-rich repeats (LRR) (Wu et al., 2016) (Fig. 2B), the malectin homologous sequences, perhaps via interaction with the cell wall, might provide a scaffold function thereby stabilizing the LRR or Kelch motif-mediated interactions with other cell surface molecules. Tandem malectin domains in plant proteins might augment the affinity for a specific target or each might interact with distinct targets, improving specificity or extending their functional range.

With the widespread presence of malectin homologous sequences across the biological kingdoms, it is interesting to consider how they might have arisen and what drove their diversification. Given the various domain compositions, the most plausible scenario could be that a single domain malectin protein had existed as a common progenitor. In all M/MLD proteins, the malectin domains are homologous, and therefore decorate the cell wall-cell membrane interface. Little is known about the MLD-RLPs, but lacking a substantial cytoplasmic domain (Fig. 2B), they might function in partnership with other RLKs to mediate downstream signaling (Jeong et al., 1999; Lee et al., 2012). Two proteins, the cytoplasmic Malectin Domain Kinesin1 and 2 (MDKIN1, 2) (Galindo-Trigo et al., 2020b), constitute an outgroup among M/MLD proteins. Why a protein domain that has largely evolved to be in the extracellular milieu or the ER lumen had become cytoplasmic and associated with a cytoskeleton protein is intriguing. MDKIN2 belongs to an angiosperm-specific clade while MDKIN1 is present in angiosperms, gymnosperms, *Selaginella selmo*, *Physcomitrella patens* and *Marchantia polymorpha*. MDKIN2 has been localized to the phragmoplast midzone during cell division where cell wall materials are being laid down while the adjacent cell membrane is not yet fully defined (Smertenko, 2018). How this characteristic might contribute to its function, characterized as non-essential but important for pollen and seed development (Galindo-Trigo et al., 2020b), will be interesting to explore.

**The Arabidopsis M/MLD-RLKs**

The M/MLD-RLKs refer collectively to the LRR–MD–RLKs, the FER/THK1 family of M/MLD-RLKs, and the MLD–LRR–RLKs (Fig. 2B).

**The LRR–MD–RLK clade.** The LRR–MD–RLKs have a LRR domain followed by a single malectin domain (Fig. 2B). These malectin domains share the highest identity with *Xenopus* malectin than those in the other clades. Several LRR–MD–RLK genes, some of which are BSRs, are transcriptionally responsive to BR (Xu et al., 2013) and could be involved in the signaling pathways regulated by this hormone. There are considerable cross-talks between BR and immunity signaling (Lozano-Díran and Zipfel, 2014). The expression of several BSRs, including *RECEPTOR-LIKE KINASE IN FLOWERS 1* (*RKF1*), *BSR430*, *BSR840/LIK1* and *A1gS56120*, was triggered during immunity response (Qutob et al., 2006, Hok et al., 2011). Interestingly, BSR840/LIK1 interacts with the LysM motif-containing chitin receptor CERK1 and loss-of-LIK1 resulted in defective jasmonic acid and ethylene signaling; mutant plants were also more resistant to bacterial and fungal infections (Le et al., 2014). These M/MLD-RLKs might bridge BR and immunity signaling cross-talks.

**The MLD–LRR clade.** This comprises a family of 17 related RLKs (Fig. 2A) that is originally referred to as the CrRLK1-L RLKs (for *Catharanthus roseus* RLK1-like RLKs), after the founding member (Schulze-Münti et al., 1996), for which no further studies have been reported. To date, the biological roles of all but one (At5g24010) of these MLD-RLKs have been examined, some in detail. Together they regulate diverse processes throughout the plant life cycle, controlling growth and mediating survival (Galindo-Trigo et al., 2016; Doblas et al., 2018; Frank et al., 2019). The fact that achieving this level of functional revelation has taken just over a decade is quite remarkable. It also suggests that ML domains, with tandem malectin homologous regions, must have provided an important functional advantage, rendering phenotypic manifestation from their loss readily notable.

The first functionally studied members of this clade, FER and THK1,
Fig. 2. M/MLD proteins in Arabidopsis. (A) Mid-point rooted phylogenetic tree of M/MLD family proteins. These proteins were first identified by searching the Arabidopsis proteome with Pfam profile (PF1721 and PF2189) using HMMER (http://hmmer.org/). Retrieved proteins were used as blastp queries against the A. thaliana proteome (E value < 1e-5). Protein domain analysis was used to verify the presence of single Malectin domains (MD) and tandem Malectin domains (MLD). Four proteins, encoded by Atg606000, Atg606280, Atg561840, Atg59616 and indicated by asterisks, either have low supporting values or have truncated ML sequences. They were not included in plant M/MLD-RLKs phylogenetic analysis. Domain organization of M/MLD family proteins. (B) The most prevalent organization for each clade seen in (A) is shown in corresponding color blocks here. MLD-RLPs lack a cytoplasmic domain and are further divided into two classes, some having a predicted transmembrane helix others LRR domains of various lengths. (C) Structural alignment of the MLD in FER with animal malectin. (Top) The FER ECD structure (pdb 6AS5, from Xio et al., 2019) with the tandem malectin domains MalA and MalB shown in red and green, respectively. (Bottom) The Xenopus Malectin structure [pdb 2JWP, from Schallus et al., 2008] is shown in blue and superimposed onto MalA and MalB. Comparable superimposition was shown with the structurally determined MLDS of ANXs (Moussu et al., 2018) and could be achieved in structures of other MLD-RLKs predicted by computational modeling (not shown). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

*H. Yang et al.*
The MLD-LRR-RLK clade. The ML domain in the MLD-LRR-RLKs are followed by a region of LRRs (Fig. 2B) and they represent a majority of the MLD-containing RLKs (Fig. 2A). Several are involved in immunity signaling, such as IOS1 (Hok et al., 2011; Yeh et al., 2016), STRESS-INDUCED FACTOR2 (SIF2) (Chan et al., 2020), FLG22-INDUCED RECEPTOR-LIKE1 (FRK1)/SENESCENCE-INDUCED RECEPTOR KINASE (SIRK1) (Yeh et al., 2015; Robatzek and Somssich, 2002), and OUTGROWTH-ASSOCIATED PROTEIN KINASE (OAK) (Smith et al. 2011). Some are associated with developmental processes, such as MATERNAL EFFECT EMBRYO ARREST39 (MEE39) and ROOT HAIR SPECIFIC (RHS) 6 and RHS16 for endosperm (Pagnussat et al., 2004) and root hairs (Won et al., 2009), respectively.

Fig. 3. Chromosomal and synteny map of ML proteins in A. thaliana. The Arabidopsis synteny information was obtained from The Samuel Robert Noble Foundation (http://plantgrn.noble.org/LegumeIPv2/download.jsp). The genetic features of M/MLD protein genes was downloaded from the Arabidopsis Information Resource (TAIR) (https://www.arabidopsis.org/portals/genAnnotation/gene_structural_annotation/agicomplete.jsp) on www.arabidopsis.org. The M/MLD genes were mapped to synteny regions. Only syntenic blocks with synteny scores higher than 1000 containing M/MLD family proteins are retained here. Pairs of syntenic regions are represented as the same color-coded blocks; syntenic gene pairs are indicated by the same color-coded arrows. Clade names are shown in blue. *, designated as indicated in Fig. 2A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Analysis of transcriptomic data extracted from Genevestigator (Hruz et al., 2008) shows that the MLD-LRR-RLK clade stands out with many of its members showing higher than a two-fold increase in expression during pathogen or elicitor treatments (Group 1, S. Fig. 1A; Group 3, S. Fig. 1B). These include IOS1, FRK1/SIRK1, and SIF2, consistent with their already elucidated roles in plant defense, and several others in the same chromosomal cluster between At1g51790 to At1g51890 (Fig. 2A; Fig. 3). These data also show that biotic stresses down-regulate a group from all three MLD-RLK clades (Group 2, S. Fig. 1A; Group 4, Suppl. Fig. 1B). To what extent these MLD-RLKs have expanded to provide functional redundancy to safeguard survival and how many of them have evolved to play specific roles under different conditions will be interesting functional and evolutionary questions to explore.

Chromosomal distribution of M/MLD protein genes in Arabidopsis

Chromosomal mapping and synteny analysis of genes encoding A. thaliana M/MLD proteins shows that many of them are tandemly arranged within segmentally duplicated regions of the genome (Fig. 3). Typically, a protein family arose from multiple duplications of an ancestral gene through whole genome duplication (WGD), gene duplication, exon duplication and shuffling (Cannon et al., 2004; França et al., 2012; Panchy et al., 2016). The A. thaliana genome has experienced WGD, gene loss, local gene duplication, transposition, and segmental duplication events, thus duplicated genes are shuffled throughout the genome (Arabidopsis Genome Initiative, 2000). The occurrence of syntenic blocks of M/MLD protein genes and local duplications is consistent with the notion that gene expansions that led to the Arabidopsis RLK superfamily had occurred via recent duplications.

Eight syntenic blocks containing M/MLD protein genes were identified (paired color-coded blocks, Fig. 3). Genes for the LRR–RLK and MLD–LRR-RLK clades display considerable duplications and local expansions. 12 of the 13 LRR–MLD–RLK genes are located on chromosome 1. BSR840 is maintained as a single gene on chromosome 3 and paired with the homolog At1g53420 in the BSR cluster on chromosome 1 (block 6). Duplication events at the At1g53420 locus likely happened giving rise to the local cluster with BSR430 and BSR440, and also dispersing BSR650 to block 1 on the same chromosome (Fig. 3). This genomic arrangement suggests an initial duplication event had given rise to these BSR blocks, which was followed by more local gene expansions.

The MLD–LRR–RLK genes also occur in multiple clusters. Pairing of the MLD–LRR–RLK cluster on chromosome 1 with the single At3g21340 on chromosome 3 (block 5) suggests an early duplication event could have generated the two loci, followed by extensive local duplications at the chromosome 1 locus to yield the large contiguous cluster. Similarly, a duplication followed by local expansion could have generated the tandem arrays of genes in the three other paired chromosomal IOS1 gene segments (blocks 1, 7, and 12). An analysis of the gene cluster between At5g59650 and At5g59680 in block 12 on chromosome 5 in more than three hundred accessions of A. thaliana and its close relative A. lyrata (Smith et al., 2011) suggests that the A. thaliana cluster had arisen through a duplication of the central RLK gene of an ancestral three-gene cluster with At5g59650 and At5g59680 flanking At5g59670 (OAK1), giving rise to At5g59660, making up the 4-gene cluster (Fig. 3). It is interesting to note that the gene for the MLD-RLK ANJ is just adjacent to the OAK1 MLD-LRR-RLK cluster but ANJ lacks the LRR-coding sequence. ANJ is paired with HERK1 in block 12 on chromosome 3, where local expression has given rise to the adjacent IOS1 cluster. An initial duplication may have happened in one form, either with or without LRR, followed by losing or gaining the LRR, respectively. A duplication of the MLD-LRR-RLK gene pair in block 1 on chromosome 1, At1g07550 and At1g07560, could have generated the block 1 MLD–LRR–RLK cluster on chromosome 3.

In contrast, most of the MLD–RLK clade genes, in particular those associated with notable biological roles, are mapped in unique chromosomal locations. Even when existing as duplicated genes, e.g. the pollen-specific ANX1 and 2 genes and BUPS1 and 2 genes, they did not further expand locally. On the other hand, the functionally subtle MDS1-4 (At5g38990-At5g39030) under normal growth conditions (Richter et al., 2018), are clearly a result of local expansion. Do the chromosomal organization of the MLD–RLK clade genes with crucial biological roles and the mostly clustered organization of the MLD–LRR–RLK, LRR–MD–RLK clade genes and MDS1-4 reflect a correlation with the biological roles maintained for these genes? Perhaps local gene expansion could be a mean to dilute functional contribution from a single progenitor with an essential function so the biological burden is shared by gene redundancy. Alternatively, it could also be a mean to augment functional impacts via expansion of a progenitor gene in a chromosomal region of low expression activities. Stability at chromosomal locations for the MLD–RLK clade gene locations could allow these genes to be preserved for conserved functional roles whereas instability at the MLD–LRR–RLK clade gene locations could have facilitated their neo-functionalization to meet varying environmental demands (Duarte et al., 2006).

M/MLD–RLKs among diverse plant genomes

Moss (Physcomitrella patens), Medicago (Medicago truncatula), tomato (Solanum lycopersicum) and rice (Oryza sativa) were selected as a representative non-vascular land plant, a model legume, a dicot and a monocot crop, respectively, for a broader phylogenetic analysis of M/MLD–RLKs. We identified 279 M/MLD-RLKs from these plants, with 87, 75, 65, 32 and 20 from Medicago, A. thaliana, rice, tomato and moss, respectively. Similar to the Arabidopsis proteins, these M/MLD–RLKs are classified into two major phylogenetic branches, the single malacatin domain-containing LRR–MLD–RLK clade and the tandem malacatin domain-containing MLD clade that further divides into the MLD–RLK and MLD–LRR–RLK subclades (Fig. 4). Each of these three clades are further divided into two subclades (Fig. 4, inner circle).

The MLD–LRR–RLK and MLD–RLK clades. These clades share evolutionary patterns with their counterparts from Arabidopsis with the LRR–MD–RLK clade displaying extensive lineage-specific local expansion while the MLD-RLK clade shows only limited expansion. Of the LRR–MLD–RLK clade, Medicago and rice genes have evolved their own lineage, arisen from local expansion after segmental duplications. Most Arabidopsis LRR–MLD–RLK clade genes do not have close relatives in these species, except for two of its members (blue arrowheads, Fig. 4).

Many close homologs of the Arabidopsis MLD–RLK clade genes are found in all four species examined. For example, HERK2 are in a group with two rice, one Medicago and one tomato genes, suggesting that they are descended from a common ancestor. Strikingly, a large family of Medicago MLD–RLKs and the Arabidopsis FER and ANXs evolved within two monophyletic clades, which are also occupied by several rice and tomato homologs. Together with the already recognized functional importance of FER- and THE1-related Arabidopsis MLD–RLKs and several FER-Like Receptors (FLRs) in rice, such as DWF4 AND RUNTISH SPIKELET1, 2 (DRUS1/FLR2 and DRUS2) in reproduction and immunity (Pu et al., 2017; Yang et al., 2020; Huang et al., 2020), it appears that genes in the MLD-RLK clade had evolved under considerable constrain for functional conservation.

The MLD–LRR–RLK clade and Symbiosis Receptor Kinases (SymRiks). The MLD–LRR–RLK subclade 1 includes a large majority of its members and comprises five lineage-specific monophyletic groups (Fig. 4). Lineages I to IV include the majority of IOS1 clade genes from Arabidopsis and the other species examined, with no or few interruptions. Lineage V has a small number of Arabidopsis, Medicago, rice and tomato genes interrupting a large moss gene cluster. Like in Arabidopsis MLD–LRR–RLK clade, the subclade1 genes from these three species had also experienced duplications and local expansions.

The Sub-MLD-LRR-RKL clade 2, with only six members, comprises...
the Symbiosis RKs (SymRKs) (Markmann et al., 2008) (Fig. 4), including the **DO NOT MAKE INFECTION2** (DMI2) from Medicago (Endre et al., 2002). In most land plants, SymRKs are known to play important roles in mediating symbiotic interactions with arbuscular mycorrhizal (AM) fungi and, in addition, nitrogen-fixing symbiosis with certain bacteria, mostly in legumes (Stracke et al., 2002; Markmann et al., 2008; Gherbi et al., 2008). Defects in SymRK abolish AM symbiosis and, where applicable, nitrogen-fixing symbiosis. Overexpressing full-length or just the kinase domain of SymRK leads to spontaneous nodule formation in the absence of nitrogen-fixing rhizobia (Markmann et al., 2008; Saha et al., 2014). A precise functional role for the ML domain in SymRK remains unclear. The rice SymRK is missing in the phylogeny map.
perceived to a hub capable of dispatching them to diverse intracellular processes (e.g., Boisson-Dernier et al., 2015; Du et al., 2016; Huang et al., 2020), the Arabidopsis MLD–RLKs studied thus far nevertheless display considerable conservation in their core signaling strategies. For example, they function as homo- and/or hetero-oligomers and interact with a glycosylphosphatidylinositol-anchored protein (GPI-AP) from the LORELEI and LLG1,2,3 protein family (Li et al., 2015; Ge et al., 2019a, 2019b; Feng et al., 2019; Galindo-Trigo et al., 2020a; Liu et al., 2021; Noble et al., 2021). Thus far, Rapid Alkalization Factor (RALF) peptides (Blackburn et al., 2020) have been identified as ligands for FER, THE1, ANXS, BUPSs and ANJ (Haruta et al., 2014; Stegmann et al., 2017; Gonneau et al., 2018; Liu et al., 2021). The FER ECD–RALF23 (an N-terminal 1–17 amino acid fragment) and LLG2, which is specific to pollen where FER is not expressed, form a stable complex in vitro (Xiao et al., 2019), suggesting that functionally non-cognate components are structurally compatible. This is consistent with the LORELEI/LLG proteins can functionally substitute for each other (Noble et al., 2021). The LORELEI/LLG proteins also chaperone the partnered receptor kinases to their functional location on the cell membrane (Li et al., 2015; Feng et al., 2019). It would be interesting to examine whether similar strategies involving variants of the MLD–RLK and LORELEI/LLG coreceptor partners could be employed by a broader set of M/MLD-RLKs and that RALFs and other unrelated peptides (Liu et al., 2021) could act as ligands to regulate their activity.

On the transmembrane and inner cell membrane levels, FER and THE1 interact with guanine nucleotide exchange factors (GEFs) to activate RHO GTPases (Duan et al., 2010, 2014; Li et al., 2015; Huang et al., 2013; Qu et al., 2017; Lin et al., 2018), thus linking signals perceived to a hub capable of dispatching them to diverse intracellular processes (Nibau and Cheung, 2011). FER and ANX are functionally linked to the RHO-regulated NADPH oxidases (Duan et al., 2010; Boisson-Dernier et al., 2013; Li et al., 2015; Liu et al., 2021; Zhang et al., 2021; Song et al., 2021), which control ROS production. ROS are ubiquitous signaling molecules whose oxidative power also impacts cell wall quality. FER, ANX1 and THE1 also interact with several other cell surface receptors important for growth, immunity and other stress signaling pathways (Stegmann et al., 2017; Dressano et al., 2017; Mang et al., 2017; Van der Does et al., 2017). Interactions with cytoplasmic receptor-like kinases (Boisson-Dernier et al., 2015; Du et al., 2016; Shen et al., 2017) and various molecular conglomerates of intracellular signaling molecules along the cytoplasmic membrane surface (Mang et al., 2017; Liu et al., 2020; Huang et al., 2020) have also been demonstrated. Additional mechanisms will inevitably emerge to further elaborate the molecular interaction networks engaged by MLD-RLKs as more are being examined.

Interacting physically and/or functionally with the cell wall, the potential property that first brought intense attention to the MLD-RLKs could also be a shared mechanistic theme. The identification of THE1 suggested that MLD-RLKs could serve as surveyors of cell wall quality to act on demand to signal the needed cellular responses (Hematy et al., 2007). FER physically interacts with homoglactururonan, the backbone of the cell wall polysaccharide pectin (Feng et al., 2018; Lin et al., 2018; Duan et al., 2020). Altered cell wall quality could contribute to growth and immunity related phenotypes in loss of function fer mutants (Doblas et al., 2018; Frank et al., 2018), in particular in root growth responses under high salinity (Feng et al., 2018) and FER-controlled process to suppress supernumerary pollen tube penetration of ovules (Duan et al., 2020). Moreover, fer mutants also lack the elaborate puzzle-like epidermal cell pattern in leaves (Li et al., 2015; Noble et al., 2021) where differential pectin deposition is believed to underlie the morphogenesis of the interdigitated pavement cells (Haas et al., 2020). ANXs or BUPSs are crucial for pollen tube wall quality that prevents precocious pollen tube rupture (Boisson-Dernier et al., 2009, 2013; Miyazaki et al., 2009; Ge et al., 2017, 2019a, 2019b). ERU affects pectin deposition in root hair cell walls and eru mutant root hairs are short and stocky (Schoenaers et al., 2018). However, although some structural homology exists (Fig. 2C), a sugar-binding pocket analogous to that in the diglucose binding animal malectin (Schallus et al., 2008) could not be assigned in the crystal structures elucidated for ANX1, ANX2 and FER (Moussu et al., 2018; Du et al., 2018; Xiao et al., 2019). At the cell membrane-wall interface, MLD-RLKs are in the vicinity of many components of the extracellular matrix in addition to peptide ligands, pectin and the LLG family proteins. These broader interactions can conceivably provide a signaling linkage to molecules that lack cytoplasmic signaling capacity, such as additional GPI-AP proteins (Hou et al., 2016) or as in the demonstrated interaction between FER and the tightly cell wall-associated Leucine-rich repeat extensins (LRXs) (Zhao et al., 2018; Dunser et al., 2019; Herger et al., 2020), linking extracellular sensing to intracellular processes.

The abundance of M/MLD–containing proteins in plants might represent a unique niche that has facilitated species diversification and survival. Well-characterized model and crop plant species are well suited to further decipher how the malectin homologous sequences in M/MLD-RLKs enable plants to respond to endogenous demands and environmental conditions in order to thrive and propagate. Understanding the functional involvement of M/MLD proteins across the broader plant kingdom should reveal interesting specialization strategies to meet the demands from diverse plant forms and unique environmental challenges faced by different species in their habitats.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

HY, LG thank Dr. Li-Jun Ma for support and guidance; HP was supported by NSF-1557994 to DW; RY was supported by (NSF-1715764) to HMW and AYC.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tcsw.2021.100056.

References

Adams, J., Kelso, R., Cooley, L., 2000. The kelch repeat superfamily of proteins: propellers of cell function. Trends Cell Biol. 10, 17–24.
Antolin-Llovera, M., Petutsching, E.K., Reid, M.K., Lipka, V., Nurnberger, T., Robatzek, S., Paradies, M., 2014. Knowing your friends and foes—plant receptor-like kinases as initiators of symbiosis or defence. New Phyto. 204, 791–802.
Bai, L., Ma, X., Zhang, G., Song, S., Zhou, Y., Gao, L., Mao, Y., Song, C.-P., 2014. A receptor-like kinase mediates ammonium homeostasis and is important for the polar growth of root hairs in Arabidopsis. Plant Cell 26, 1497–1511.
Boisson-Dernier, A., Parniske, M., 2014. Knowing your friends and foes—plant receptor-like kinases as initiators of symbiosis or defence. New Phyto. 204, 791–802.
Blackburn, M.R., Haruta, M., Duan, Q.-H., 2020. Twenty Years of Progress in Physiological and Biochemical Investigation of RALF Peptides. Plant Physiol. 182, 1065–1066.

Boisson-Dernier, A., Frank, C.M., Litvüe, D.S., Grossniklaus, U., 2015. Receptor-like cytoplasmic kinase MARIS functions downstream of CRKLI-LK-dependent signaling during tip growth. Proc. Natl. Acad. Sci. U.S.A. 112, 12211–12216.

Boisson-Dernier, A., Kita, D., Peaucelle, A., Wu, H.-M., 2019. Cell wall integrity signaling in plants: malectin domain receptors. Annu. Rev. Plant Biol. 70, 1581–1591.

Boisson-Dernier, A., Litvüe, D.S., Nestorova, A., Frank, C.M., Thirumaranarajah, S., Grossniklaus, U., 2013. ANXUR receptor-like kinase coordinates cell wall integrity with growth at the pollen tube tip via NADPH oxidases. Plos Biol. 11, e1001719.

Boisson-Dernier, A., Roy, S., Kritsas, K., Grobei, M.A., Jaciubek, M., Schroeder, J.I., 2016. Identification of a receptor-like kinase, related to cell wall expansion during cellular elongation in Arabidopsis thaliana. EMBO J. 35, 3265–3279.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., May, G., 2004. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Biol. 2, 17.

Duan, Q.-H., Kita, D., Johnson, E.A., Aggarwal, M., Gates, L., Wu, H.-M., Cheung, A.Y., Feng, W., Kita, D., Peaucelle, A., Doan, V., Duan, Q.-H., Liu, M.-C., Maman, J., Yvon, R., Duan, C.-J., Feng, Y.-L., Cao, Q.-L., Huang, M.-Y., Feng, J.-X., 2016. Identification of a receptor-like kinase, related to cell wall expansion during cellular elongation in Arabidopsis thaliana. BMC Biol. 2, 17.

Gachomo, E.W., Jon Baptist, L., Kefela, T., Saidel, W.M., Kotcheni, S.O., 2014. The Arabidopsis CURVY (CVY1) gene encoding a novel receptor-like protein kinase regulates cell morphogenesis and promotes seed recruitment and seed production. BMC Plant Biol. 14, 218.

Galindo-Trigo, S., Gray, J.E., SMITH, L.M., 2016. Conserved roles of CRKLI receptor-like kinases in cell expansion and reproduction from algae to angiosperms. Front. Plant Sci. 7, 1606.

Galindo-Trigo, S., Blanco-Tourinan, N., DeFalco, T.A., Wells, E.S., Gray, J.E., Zipfel, C., Smith, L.M., 2020a. CRKLI receptor-like kinase HERK1 and ANJEA are female determinants of pollen tube release. EMBO Rep. 21, e48194.

Galindo-Trigo, S., Grand, T.M., Voigt, C.A., SMITH, L.M., 2020b. A malectin domain kinase functions in pollen and seed development in Arabidopsis. J. Exp. Bot. 71, 1826–1841.

Galli, V., Pascasiano, R., Solda, T., Calanca, V., Molinari, M., 2011. Malectin participates in a backup glycoprotein quality control pathway in the mammalian ER. 2011. Plos One 6, e16304.
Takeda, K., Qin, S.-Y., Matsumoto, N., Yamamoto, K., 2014. Association of malectin with ribophorin 1 is crucial for attenuation of misfolded glycoprotein secretion. BBRC 454, 436–440.

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gantzbiel, L., Childs, K.L., Yandell, M., Gundlach, H., Mayer, K.F., Schwartz, D.C., Town, C.D., 2014. An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC genomics 15, 312.

Tannous, A., Pisoni, G.B., Hebert, D.N., Molinari, M., 2014. N-linked sugar-regulated protein folding and quality control in the ER. Semin. Cell Dev. Biol. 41, 79–89.

The Arabidopsis Genome Initiative, The Arabidopsis Genome, 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815.

Van der Does, Dioneurjo, Boutrot, Freddy, Engelisdorf, Timo, Rhodes, Jack, McKenna, Joseph F., Vernhnette, Samantha, Krooavelts, Iku, Tintor, Nico, Veerabagu, Manikandan, Miedes, Eva, Segonzac, Cécile, Roux, Mileina, Breda, Alice S., Hardke, Christian S., Molina, Antonio, Rep, Martin, Testerink, Christa, Mouillé, Gregory, Hofte, Herman, Hamann, Thorsten, Zipfel, Cyril, Cheung, Alice, 2017. The Arabidopsis leucine-rich repeat receptor kinase MRK2/LRR-KISS connects cell wall integrity sensing, root growth and response to abiotic and biotic stresses. Plos Gen. 13 (6), e1006832. https://doi.org/10.1371/journal.pgen.1006832.

Wang, G., Ellendorff, U., Kemp, B., Mansfield, J.W., Forsyth, A., Mitchell, K., Bastas, K., Liu, C.-M., Woods-Tor, A., Zipfel, C., de Wit, M.J.G., Jones, J.D.G., Tor, M., Thomma, B.P.H.J., 2008. A genome-wide functional investigation into the roles of receptor-like proteins in Arabidopsis. Plant Physiol. 147, 503–517.

Wang, G., Fiers, M., Ellendorff, U., Wang, Z., Pierre, J.G., de Wit, M., Angenent, G.C., Thomma, B.P.H.J., 2010. The diverse roles of extracellular leucine-rich repeat-containing receptor-like proteins in Plants. Crit. Rev. Plant Sci. 29, 285–299.

Won, S.-K., Lee, Y.-J., Lee, H.-Y., Cha, K.-Y., Cho, M., Cho, H.-T., 2009. Cis-element- and transcriptome-based screening of root hair-specific genes and their functional characterization in Arabidopsis. Plant Physiol. 150, 1459–1473.

Yu, F., Qian, L., Nibau, C., Duan, Q., Li, X., Levansus, K., Li, X., Li, H., Hou, C., Li, L., Buchanan, B.B., Cheung, A.Y., Li, D., Luan, S., 2012. FERONIA receptor kinase pathway suppresses abscisic acid signaling in Arabidopsis by activating ABI2 phosphatase. Proc. Natl. Acad. Sci. U.S.A. 109, 14694–14698.

Zhang, L., Huang, J., Su, S., Wei, X., Yang, L., Zhao, H., Yu, Y., Wang, J., Hui, J., Hao, S., Song, S., Cao, Y., Wang, M., Zhang, X., Zhao, Y., Wang, Z., Zeng, W., Wu, H.-M., Yuan, Y., Cheung, A.Y., Duan, Q., 2021. FERONIA-regulated reactive oxygen species mediate self-incompatibility in Brassica rapa. Current Biology, online. https://doi.org/10.1016/j.cub.2021.04.060.

C. Zhao O. Zayed Z. Yu W. Jiang P. Zhu C.-C. Hsu L. Zhang W.A. Tao R. Lozano-Duran J.-K. Zhu Leucine-rich repeat extension proteins regulate plant salt tolerance in Arabidopsis 2018 Natl. Acad. Sci. U.S.A. www.pnas.org/cgi/doi/ Proc 10.1073/ pnas.1816991115.