doi:10.1093/plcell/koab173

THE PLANT CELL

THE PLANT CELL 2021: 33: 2915–2934

Perspective

The sequenced genomes of nonflowering land plants reveal the innovative evolutionary history of peptide signaling

Chihiro Furumizu [1,†], Anders K. Krabberød [2], Marta Hammerstad [3], Renate M. Alling [2], Mari Wildhagen [2], Shinichiro Sawa [1] and Reidunn B. Aalen [2,*,‡]

1 Graduate School of Science and Technology, Kumamoto University, 860-8555 Kumamoto, Japan
2 Section for Evolutionary Biology and Genetics, Department of Biosciences, University of Oslo, 0316 Oslo, Norway
3 Section for Biochemistry and Molecular Biology, Department of Biosciences, University of Oslo, 0316 Oslo, Norway

*Author for correspondence: reidunn.aalen@ibv.uio.no
†Present address: Department of Immunology, Graduate School of Medicine, Akita University, 010-8543 Akita, Japan.
‡Senior author.

M.W. and S.S. initiated and drafted the work; C.F., A.K.K., R.M.A., and R.B.A. identified and analyzed sequences; M.H. performed structure analyses and modeling; C.F. and R.B.A. wrote the article. C.F. created Figures 1, 6, and 8, and Table 1 with supporting tables and figures with input from R.M.A. and R.B.A.; A.K.K. generated Figures 2–4, and Tables 2 and 3 with supporting figures with input from R.M.A. and R.B.A.; Figures 5 and 7 with supporting figures were created by M.H. and R.B.A.

The author(s) responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell) are: Reidunn B. Aalen (reidunn.aalen@ibv.uio.no) and Chihiro Furumizu (chihiro.furumizu@med.akita-u.ac.jp).

Abstract

An understanding of land plant evolution is a prerequisite for in-depth knowledge of plant biology. Here we extract and explore information hidden in the increasing number of sequenced plant genomes, from bryophytes to angiosperms, to elucidate a specific biological question—how peptide signaling evolved. To conquer land and cope with changing environmental conditions, plants have gone through transformations that must have required innovations in cell-to-cell communication. We discuss peptides mediating endogenous and exogenous changes by interaction with receptors activating intracellular molecular signaling. Signaling peptides were discovered in angiosperms and operate in tissues and organs such as flowers, seeds, vasculature, and 3D meristems that are not universally conserved across land plants. Nevertheless, orthologs of angiosperm peptides and receptors have been identified in nonangiosperms. These discoveries provoke questions regarding coevolution of ligands and their receptors, and whether de novo interactions in peptide signaling pathways may have contributed to generate novel traits in land plants. The answers to such questions will have profound implications for the understanding of the evolution of cell-to-cell communication and the wealth of diversified terrestrial plants. Under this perspective, we have generated, analyzed, and reviewed phylogenetic, genomic, structural, and functional data to elucidate the evolution of peptide signaling.

Received January 20, 2021. Accepted June 08, 2021. Advance access publication July 7, 2021

Introduction

Embryophytes evolved from freshwater green algae that began to colonize land approximately 470 million years ago (Mya; Delwiche and Cooper, 2015; Ishizaki, 2017). This evolutionary shift required key biological innovations, including the transition from the ancestral algal life cycle of haploid multicellular plants with zygotic meiosis, to alternating multicellular haploid (gametophyte) and diploid (sporophyte) generations. Although the diploid sporophyte and the haploid gametophyte initially might have employed the same genes for growth and development, gene duplications followed by neofunctionalization likely facilitated the development of new specialized organs by extensive changes in gene expression patterns (Ligrone et al., 2012; Bowman et al., 2017; Jill Harrison, 2017; Figure 1). The evolution of complex vasculature in ancestors of seed plants, lycophytes, and ferns, which appeared around 430 Mya (Kenrick and Crane, 1997; Gensel, 2008; Steemans et al., 2009; Morris et al., 2018), provided mechanical support and enabled improved transport of nutrients (Ishizaki, 2017). Roots evolved independently in lycophytes and euphyllophytes (vascular plants excepting lycophytes) providing anchorage for the plants and enabling acquisition of water and nutrients from the soil (Raven and Edwards, 2001; Hetherington and Dolan, 2019). The evolution of novel cell types, tissue types, and organs with increasing multicellular complexity likely required tuning of cell-to-cell communication in early land plants (Grienenberger and Fletcher, 2015).

Recent advances in whole-genome sequencing of land plants, as well as streptophyte algae, have facilitated large-scale comparative genomic studies. Such studies are helping us to understand the evolutionary processes of plant terrestrialization and diversification (Rensing et al., 2008; Banks et al., 2011; Amborella, 2013; Nystedt et al., 2013; Bowman et al., 2017; Li et al., 2018; Nishiyama et al., 2018; Cheng et al., 2019; Chen et al., 2020; Jiao et al., 2020; Li et al., 2020; Yang et al., 2020; Zhang et al., 2020a). For instance, the ancestry of some components of hormone and stress signaling pathways in land plants has been traced back to algae (Bowman et al., 2017; de Vries et al., 2018). Similarly, in flowering

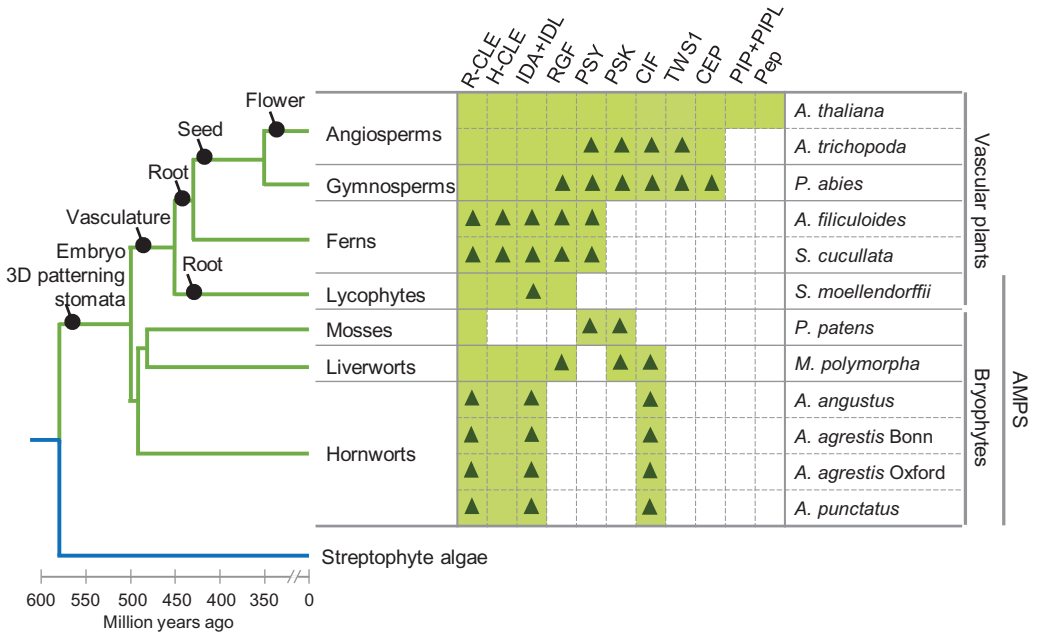

Figure 1 Phylogenetic distribution of small PTMPs in land plants. Left: Simplified phylogenetic tree of major extant land plant lineages (green lines) based on (One Thousand Plant Transcriptomes Initiative, 2019). The branch lengths are roughly proportional to the estimated divergence dates (Morris et al., 2018). Gains of key morphological innovations are shown as black circles mapped on the tree while some gain events are still under discussion. Right: Green-colored boxes indicate the presence of homologs of the indicated small posttranslationally modified signaling peptides for the model plants *Arabidopsis thaliana* (Lamesch et al., 2012), *Amborella trichopoda* (Amborella Genome Project, 2013), *Picea abies* (Nystedt et al., 2013), *Azolla filiculoides* and *Salvinia cucullata* (Li et al., 2018), *S. moellendorffii* (Banks et al., 2011), *Physcomitrium patens* (Lang et al., 2018) *Marchantia polymorpha* (Bowman et al., 2017), and *Anthoceros* sp. (Li et al., 2020; Zhang et al., 2020a). Triangles indicate putative peptide ligands newly identified in our sequence analyses. Novel peptide-encoding genes were identified in tBLASTn searches against genomic DNA or RNA collections (e.g. *IDA* in *S. moellendorffii*), or BLASTp searches adapted for short amino acid queries using the PAM30 matrix (e.g. *CLE* genes in ferns; Altschul et al., 1997). PHI BLASTp using the aspartic acid (Asp/D) Tyrosine (Tyr/Y) motif present in RGF, CIF, PSK and PSY and/or use of wildcards in positions where the amino acid residues that are less conserved (e.g. DYxxxxxKPPIHN for RGFs), facilitated successful identification of such peptides in *P. patens*, *M. polymorpha*, ferns and gymnosperms. Genes encoding CIF and IDA peptides in the genomes of *Anthoceros* species were identified amongst coding sequences of less than 300 amino acids, using queries DYxxxxPxPPLxxPxPF and PIPxSxPSKRHN with wildcards based on MUSCLE alignments (Edgar, 2004) of previously identified CIF and IDA peptides, and the EMBOSS 6.5.7 tool fuzzpro (http://emboss.open-bio.org/rel/rel6/apps/fuzzpro.html).

plants (angiosperms), the function of classical hormones as the major mediators of signaling processes is augmented by small, secreted peptides (Matsubayashi, 2014; Olsson et al., 2019). These peptide hormones enable communication between neighboring cells through interactions with leucine-rich repeat (LRR) ectodomains of plasma membrane-bound receptors that have a cytoplasmic serine/threonine kinase domain (KD). The first small signaling peptide in plants was discovered 30 years ago (Pearce et al., 1991). Since then the increase in the number and quality of sequenced genomes, and advances in methods for in silico annotation of small genes, have together allowed the identification of hundreds of potential signaling peptides (Gong et al., 2002; Ghorbani et al., 2015; Tavormina et al., 2015; Goad et al., 2017). The peptides with known receptors comprise two major groups, the cysteine-rich peptides, which attain their 3D structure by disulfide bridges formed between pairs of cysteines, and the posttranslationally modified peptides (PTMPs; Table 1; Supplemental Table S1 and S2), generated from prepropeptides and processed to mature peptides of 5–20 amino acids (Matsubayashi, 2011). The main receptors of both peptide groups are members of a large family of LRR receptor-like kinases (RLKs), which have been assigned to 15 subfamilies according to their unique structure, organization, and number of LRRs (Shiu and Bleecker, 2001; Liu et al., 2017). Here we will focus on the ancestry of the known PTMPs, which signal almost exclusively through a subfamily with more than 20 LRRs identified in Arabidopsis as LRR-RLK XI (hereafter LRXI).

LRR-RLKs have from their discovery been studied from a phylogenetic and evolutionary perspective (Shiu and Bleecker, 2001; Lehti-Shiu et al., 2009; Liu et al., 2017; Chakraborty et al., 2019; Hosseini et al., 2020; Man et al., 2020). Signaling peptides, on the other hand, were discovered mainly due to their involvement in developmental processes, such as root growth, organ abscission, or seed development in Arabidopsis, and phylogenetic investigations have focused on their distribution and function within the orders of flowering plants (Delay et al., 2013; Lori et al., 2015; Stø et al., 2015). Subsequent studies have matched LRR-RLKs with endogenous peptides as signaling pairs in angiosperms. A number of papers have reviewed specific peptide or receptor families involved in peptide signaling (Fernandez et al., 2013a; Muschietti and Wengier, 2018; Oh et al., 2018; Taleski et al., 2018; Kaufmann and Sauter, 2019; Segonzac and Monaghan, 2019; Shi et al., 2019) and (Supplemental Table S2). However, comprehensive investigations of how widely these receptors and ligands are conserved are lacking. The increase in available genome sequences of nonflowering plants makes it feasible to elucidate peptide signaling and the likely coevolution of ligands and receptors from an evolutionary perspective.

We consider that changes in peptide signaling systems might be linked to cell-to-cell communication pathways

Table 1 Major PTMP families in land plants

Peptide Family	Major Receptor(s) (LRR-RLK subfamily)	Amino Acid Sequence of Representative Mature Peptide[a]	
		A. thaliana	*M. polymorpha*
CEP-C-TERMINALLY ENCODED PEPTIDEs	CEPR-CEP RECEPTOR (XI)	CEP1: DFROTNPGNSOGVGH	
CIF-CASPARIAN STRIP INTEGRITY FACTORs	GSO1/SGN3, GSO2, GASSHO/ SCENGEN (XI)	CIF1: DYGNNSOSORLERPPFKLIPN	Putative MpCIF: DYCDCYGFDPSPPLVHAEITF
CLE-CLAVATA3/EMBRYO SURROUNDING REGION-related	CLV1, BAM-CLAVATA1, BARELY ANY MERISTEM	CLV3:	MpCLE2:
(CLV3/R-type)	(XI)	RTVOSGoDPLHH	KEVPNGPNPLHN
CLE	PXY/TDR-PXY, TDIF RECEPTOR	CLE41:	MpCLE1:
(TDIF/H-type)	(XI)	HEVOSGONPISN	HKNPAGPNPIGN
IDA + IDL-INFLORESCENCE DEFICIENT IN ABSCISSION, IDA-LIKE	HAE, HSL-HAESA, HAE-LIKE (XI)	IDA: PIPPSAOSKRHN	MpIDA1: EVPPQGPSPIHN
Pep-plant elicitor peptides	PEPR - PEP RECEPTOR (XI)	Pep1: ATKVKAKQRGKEKVSSGRPGQHN	
PIP + PIPL-PAMP-INDUCED SECRETED PEPTIDE, PIP-LIKE	RLK7-RECEPTOR-LIKE KINASE 7 (XI)	TOLS2/PIPL3: ASGOSRRGAGH	
PSK-PHYTOSULFOKIN	PSKR-PSK RECEPTOR (X)	PSK1: YIYTQ	Putative MpPSK: YIYTSQ
PSY1-PLANT PEPTIDE CONTAINING SULFATED TYROSINE1	PSYR-PSY1 RECEPTOR (X)	PSY1: DYGDPSANPKHDPGVoOS	
RGF/GLV/CLEL -ROOT GROWTH FACTOR,GOLVEN, CLE-LIKE	RGFR/RGI-RGF RECEPTOR, ROOT GROWTH FACTOR INSENSITIVE (XI)	RGF1/GLV11/CLEL8: DYSNPGHHPORHN	Putative MpRGF: DYAEPDTHPPESN

[a]O, hydroxylated proline; o, arabinosylated hydroxyproline; and Y, sulfated tyrosine.

that facilitated the evolution from ancestral algae to land plants and revolutionized terrestrial ecosystems. To address this hypothesis, we started by perusing the literature and data on peptide signaling and LRXI. We identified the need to search for homologs of Arabidopsis *LRXI* genes in non-flowering plants as well as genes encoding specific peptides to match the presence of their presumptive receptors. Thus, our Perspective reviews published data and is also supplemented with novel findings (Supplemental Data Set S1). We offer cross-comparisons of multiple peptide ligand–receptor pairs and argue that if we are to understand the evolutionary history and innovative importance of signaling peptides in land plant evolution, we need to study cell-to-cell communication through peptide signaling pathways in an integrative manner, using diverse in silico methods that can generate hypotheses to be addressed by experimental approaches.

Evolution of subfamily XI receptors

Evolutionary constraints on LRR and KD domains

Shiu and Bleecker (2001) pioneered the characterization of Arabidopsis RLKs with ectodomains recognizing different types of ligands and highly conserved KD that have been used in previous phylogenetic analyses to study relationships among subfamilies (Zan et al., 2013; Wei et al., 2015; Liu et al., 2016; Magalhães et al., 2016; Zhou et al., 2016; Bowman et al., 2017; Lin et al., 2018; Man et al., 2020). Our focus, however, is different; namely, to better explore the evolutionary history of LRXI receptors known to interact with signaling peptides in angiosperms. Thus, for phylogenetic analysis, 27 Arabidopsis LRXI were used to sample putative LRXI homologs from 11 species including bryophytes, lycophytes, mono-, and dicotyledons, as well as *Amborella trichopoda*, a sister lineage to all other extant angiosperms. We also identified LRR-RLKs with more than 20 LRRs from the streptophyte algae (the grade of successive sister lineages to land plants) and used them as the outgroup to infer the origin of LRXI. After removal of duplicates and isoforms, and trimming, domains were identified in a final dataset consisting of 249 sequences (Supplemental Data Set S1), which were split on each side of the transmembrane region identified by InterProscan (Jones, 2014). Separate phylogenetic trees were then constructed for the ectodomain and KD segments of the *LRXI* genes (Supplemental Files S1–S5).

We noted that for most of the clades, the branches were longer in the LRR ectodomain tree compared to the KD tree. This difference suggested that the LRR evolved faster than the KD. To substantiate this, we calculated the pairwise ratio (ω) between nonsynonymous substitutions at nonsynonymous sites (K_a), and synonymous substitutions at synonymous sites (K_s), for the KDs and the LRRs encoding DNA sequences (Figure 2). The median ratio for all the angiosperm *LRXI* genes was less than 1, with a significantly lower value for KD ($\omega_{KD} = 0.015$) than for the LRR ($\omega_{LRR} = 0.024$, p-value $< 2.2 \times 10^{-16}$; Figure 2). This supports the idea that the KD is under a stronger purifying selective pressure than the LRR ectodomain. This was a common trend for the angiosperm *LRXI* genes, as ω_{KD} of 15 of the 21 subclades was significantly lower than the ω_{LRR}, according to the Wilcoxon test (significance level $P < 0.05$; Table 2).

These differences in selection pressure can be explained by the differential structural and functional requirements facing the respective domains. An overall KD configuration must be preserved to secure the kinase activity by maintaining its compact structure with 13 tightly positioned conserved motifs (Liu et al., 2017). On the other hand, the ectodomain requires a structural scaffold making up nearly half of each LRR, and in addition, amino acid residues involved in peptide recognition. The land plant LRR motif consists of 24 residues with the consensus sequence of PxxLxxLxxLxxLxxNxLxGxI (P—Pro; L—Leu; N—Asn: G—Gly; I—Ile; x—any amino acid). The LRR units that are most important in peptide binding often deviate from the strict consensus of the scaffold. Conserved nonscaffold amino acids are on the other hand likely to represent functional sites, including residues interacting with ligands or coreceptors.

Seven ancient clades of subfamily XI receptors

The phylogenetic tree based on the LRR domain, rooted by the algal outgroup LRRs, comprises seven major clades (Figure 3, A–G in LRR; Supplemental File S1 for all support values). The clades, named after the relevant Arabidopsis receptors (PEP RECEPTORs - PEPR, GASSHO - GSO, At1g17230/At5G63930/At2G33170, ROOT GROWTH FACTOR INSENSITIVE - RGI, HAESA-LIKE - HSL/CEP RECEPTOR - CEPR, TDIF RECEPTOR - TDR, and BARLEY ANY MERISTEM - BAM/CLAVATA1 - CLV1), are defined by an early branch consisting of a combination of one or more orthologs from bryophytes, and from the lycophyte *S. moellendorffii*, together with a number of receptor orthologs from *A. trichopoda*, monocots, and dicots including *A. thaliana*. All seven major clades have high support (Shimodaira–Hasegawa approximate likelihood ratio test [SH-alrt] >95; Shimodaira and Hasegawa, 1999; Supplemental File S1). However, a comparison of the LRR tree with the corresponding KD tree (Supplemental File S1) revealed differences in the phylogenetic relationship between PEPR/GSO and other groups (Figure 3, A and B in LRR and KD). The phylogenetic trees could not unambiguously resolve the relationship between the hornwort *Antheceros agrestis*, the liverwort *Marchantia polymorpha*, the moss *Physcomitrium patens*, the vascular lycophyte *Selaginella moellendorffii* (for convenience collectively named AMPS based on the first letter in their names), and the angiosperm homologs as the support values for the deep nodes were low in both phylogenies, reflecting a common problem of inferring ancient relationships for LRR-RLKs (Liu et al., 2017). The PEPR (Figure 3A) and the GSO (Figure 3B) form a paraphyletic group basally to the remaining five land plant clades. This suggests that PEPR and GSO sequences are more closely related to the algal sequences (Figure 3, A and B).

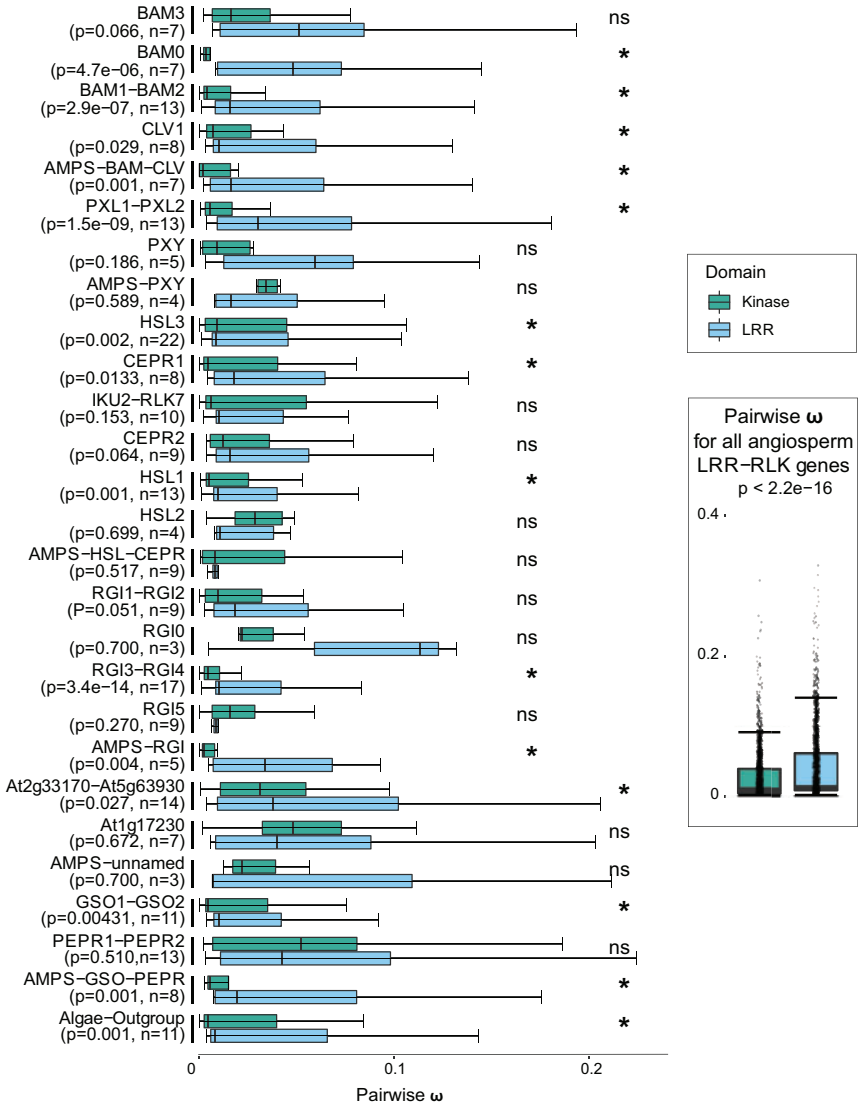

Figure 2 The LRR ectodomain evolved faster than the KD. The pairwise ratio (ω) between nonsynonymous substitutions at nonsynonymous sites (Ka) and synonymous substitutions at synonymous sites (Ks) was calculated for phylogenetically identified subclades of the clades *PEPR*, *GSO*, *Atan*, *RGI*, *HSL*, *TDR/PXY*, and *CLV/BAM* as defined from the phylogenetic tree of the LRR ectodomain (Figure 3; Supplementary File1). All subclades show purifying selection ($\omega < 1$). The boxplots delineate the median and the interquartile range of ω. The error bars extend from a value 1.5 times smaller to a value 1.5 times larger than the interquartile range. p-values were calculated based on a Wilcoxon test. *, Subclade with significant difference between ω_{LRR} and ω_{KD} ($P < 0.05$); *n*, number of sequences tested in the subclade; ns, not significant ($p > 0.05$). Insert: ω ratio between all pairs of sequences, both within and between subclades.Each pairwise comparison is marked with a point. Outliers have been removed from all boxplots for clarity.

One of the most plausible hypotheses of the land plant phylogeny proposes an early separation of both monophyletic bryophytes and vascular plant lineages (Figure 1; One Thousand Plant Transcriptomes Initiative 2019; Li et al., 2020; Zhang et al., 2020b). In line with this scenario, our phylogenetic analyses indicate that the 27 *LRXI* Arabidopsis genes originate from six or seven ancestral genes present in a common ancestor of bryophytes and vascular plants (Figure 3, Supplemental File 1and Table 3; Bowman et al., 2017; Liu et al., 2017; Li et al., 2020; Yang et al., 2020; Zhang et al., 2020a). After the two lineages split, these clades followed different trajectories. The bryophytes evolved into three distinct lineages: hornworts, mosses, and liverworts. This is reflected in their genomes, which have differentially gained and lost gene families. While *M. polymorpha* lack *PEPRs* orthologs and the clade with three *A. thaliana* genes with unknown function (hereafter named *Atanonymous*, *Atan*), the newly sequenced *Anthoceros* species also lack *Atan*, and in addition *RGIs*, while *P. patens* also lack *RGIs* and in addition *TDR* (Table 3). How these losses may have restricted the evolution of these three bryophytes remains to be elucidated.

Table 2 Median values of pairwise K_a/K_s (ω) for subclades of angiosperm LRR-RLKs subfamily XI genes with pvalues from Wilcoxcon rank-sum test

Subclade[a]	ω_{KD}	ω_{LRR}	P-Value[b]
BAM3	0.0176	0.0523	0.066
BAM0	0.0047	0.0492	**4.7×10^{-05}**
BAM1-BAM2	0.0053	0.0169	**2.9×10^{-07}**
CLV1	0.0083	0.0110	**0.029**
AMPS-BAM-CLV	0.0029	0.0174	**0.001**
PXL1-PXL2	0.0068	0.0313	**1.5×10^{-09}**
PXY	0.0207	0.0245	0.186
AMpS-PXY	0.0351	0.0173	0.589
HSL3	0.0100	0.0099	0.002
CEPR1	0.0054	0.0187	**0.013**
IKU2-RKL7	0.0070	0.0114	0.153
CEPR2	0.0135	0.0171	**0.064**
HSL1	0.0060	0.0109	**0.001**
HSL2	0.0298	0.0119	0.699
AMPS-HSL-CEPR	0.0090	0.0093	0.517
RGI1-RGI2	0.0106	0.0192	0.051
RGI0	0.0231	0.1142	0.7
RGI3-RGI4	0.0055	0.0111	**3.4×10^{-14}**
RGI5	0.0166	0.0096	0.27
aMpS-RGI	0.0033	0.0346	**0.004**
At2g33170-At563930	0.0323	0.0386	**0.027**
At1g17230	0.0490	0.0409	0.672
amPS-unnamed	0.0230	0.0081	0.7
GSO1-GSO2	0.0056	0.0112	**0.004**
PEPR1-PEPR2	0.0531	0.0436	0.51
AMPS-GSO-PEPR	0.0066	0.0203	**0.001**

[a]For convenience the three bryophytes *A. agrestis*, *M. polymorpha* and *P. patens* and the lycophyte *S. moellendorffii* are collectively named AMPS after their first letter, and a capital letter denotes presence of the relevant ortholog in the given subclade.
[b]Significantly lower ω_{KD} than ω_{LRR} in bold ($P < 0.05$); significantly higher ω_{KD} than ω_{LRR} underlined ($P < 0.05$).

In contrast, in the vascular plant lineage, *S. moellendorffii* has members in all seven clades, each with one to three paralogs, and altogether 14 *LRXIs*, indicating lineage-specific gene duplications for lycophytes (Table 3). In *A. trichopoda*, representing the sister lineage to all other flowering plants, the number of LRXI receptors has increased to 19, implicating duplications that may have contributed to the origin and early evolution of angiosperms. Nine of the present-day Arabidopsis *LRXIs* are likely results of genome duplication during the evolution of the eudicot species (Table 3; Vanneste et al., 2014). Independent gene duplications have occurred in the monocots (Figure 3).

Success stories—the *HSL* clade and the *BAM/CLV* clade

In the vascular plant genomes analyzed in this study, none of the ancestral *LRXI* genes were lost, suggesting that they all are essential for the evolution of vascular plants. In contrast, each bryophyte lineage experienced losses of different ancestral genes (Table 3). Consequently, only the two related clades, *HSL* (Figure 3E) and *CLV1/BAM3* (Figure 3G), are universally present in all lineages in our analyses, with the former showing a notable increase in gene copies in angiosperms. Gene duplications and subsequent neofunctionalization in the *HSL* clade may have refined cell-to-cell communication and facilitated the evolution of flowering plants. The *HSL* clade forms three major branches. In one branch that includes *A. thaliana HAESA* (*HAE*) and *HSL2*, *HAE* appeared during the evolution of dicots (Stø et al., 2015). These two receptors redundantly control cell separation events (Cho et al., 2008; Stenvik et al., 2008a; Shi et al., 2019). In another branch, the ancestral *HSL3* gene has proliferated both before and after the rise of flowering plants, but especially in monocots (Figure 3E; Supplemental File S1). The *HSL3* clade was the only subclade for which $\omega_{KD} > \omega_{LRR}$ ($p = 0.002$), suggesting that with the expanded number of paralogs the selective pressure was relaxed on the KD relative to LRR domain. *HSL3* was very recently shown to be involved in the regulation of stomata closure and drought stress responses by modulation of H_2O_2 levels in guard cells (Liu et al., 2020).

The last subclade includes *CEPRs* mediating nitrogen starvation signaling and controling root system architecture, (Tabata et al., 2014), and *RLK7* involved in innate immunity and lateral root formation (Hou et al., 2014; Toyokura et al., 2019), and finally, there is a recent duplication in the eudicot linage leading to *HAIKU2* (*IKU2*, Figure 3E). *IKU2* is specifically expressed during seed development and controls seed size in Arabidopsis (Luo et al., 2005) and may have been recruited to regulate angiosperm- or eudicot-specific characters (Friedman and Williams, 2004).

Interestingly, a comparison of the phylogenetic trees based on the LRR and the KD shows different branching patterns for the two *CEPR* receptors; *RLK7/IKU2* is monophyletic with *CEPR1* in the KD tree but forms a clade with *CEPR2* in the LRR tree (Figure 4A; Supplemental File S1). Closer inspection of the alignment of these four receptors revealed that CEPR1 has lost a repeat compared to the other three (Figure 4B). In the related HAE receptor, this repeat is involved in the interaction with the SERK1 coreceptor (Santiago et al., 2016). In the KD, RLK7 and IKU2 are more similar to CEPR1, and the CEPR2 is deviating, for instance, with a truncated kinase motif M3 (Figure 4C). Thus, RLK7 and IKU2 may be more like CEPR2 with regards to perceiving signals, but more similar to CEPR1 with regards to output from the KD. Altogether, the HSL clade suggests that gene duplications in the subclades, differential specificity in gene expression and, additionally, changes in the KD and the LRR influencing receptor output and ligand recognition, may contribute to evolutionary changes in biological function. Thus, the number and diversity of LRXI may have facilitated the rapid radiation of angiosperms.

Another cosmopolitan clade, *BAM/CLV1*, originated from an ancient duplication in the common ancestor of vascular plants and bryophytes, which produced its sister clade, *TDR/PXY*. Subsequently, *TDR/PXY* was lost in the moss lineage that includes *P. patens* (Figure 3F; Supplemental File S1; Table 3), illustrating differential evolutionary pressures acting on two sister clades. Within the monophyletic clade of vascular plant *TDR/PXY* sequences, two *S. moellendorffii*

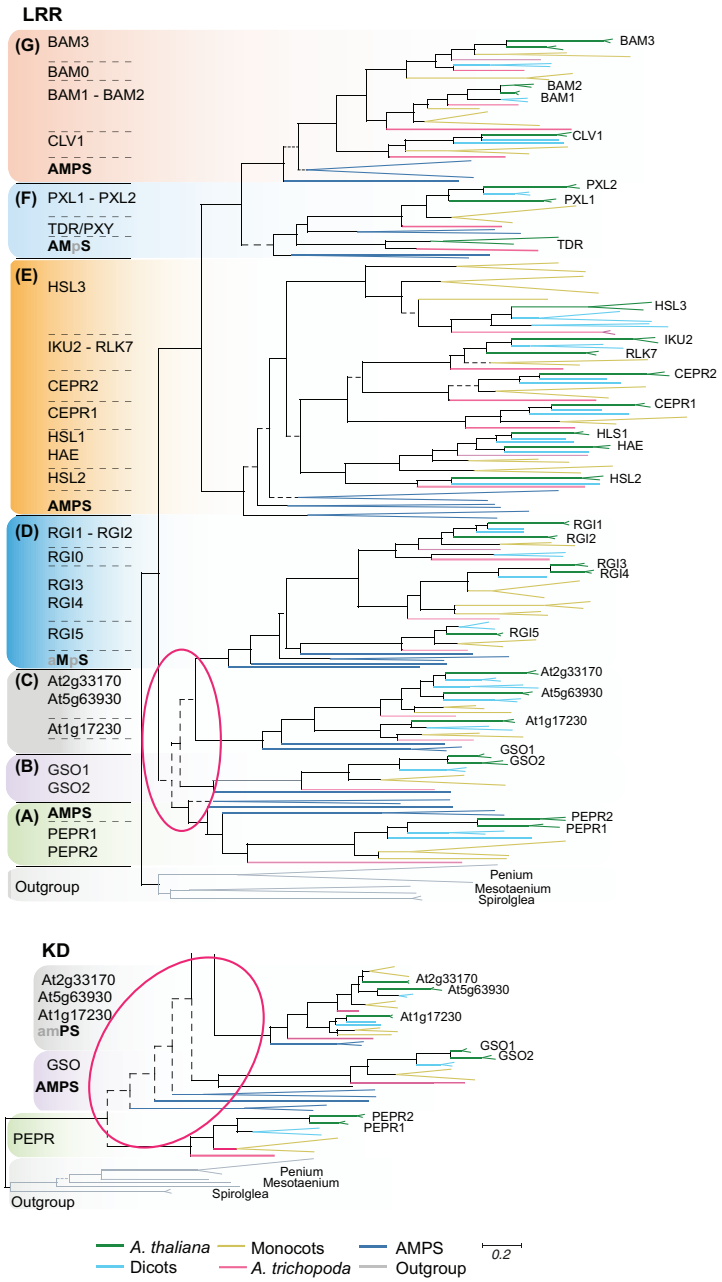

Figure 3 Phylogeny of the LRR ectodomains encoded by *LRXI* genes. LRR: Maximum likelihood phylogeny for the LRR domain encoded by 249 LRXI genes from 12 species constructed with IQ-tree (Nguyen et al., 2015) and based on a 755 amino acids long alignment, aligned with MAFFT l-INS-I and trimmed with trimal with the gappyout setting (See Supplemental File S1; Supplemental Data sets S1–S5 for details). A–G, The seven colored boxes represent monophyletic clades of the indicated angiosperm *LRXIs* with a common origin and at least one common bryophyte and lycophyte ancestor. These include (see blue branches) the hornwort *Anthoceros agrestis* (Frangedakis et al., 2021; Li et al., 2020; Zhang et al., 2020b), the liverwort *Marchantia polymorpha*, the moss *Physcomitrium patens*, and the lycophyte *S. moellendorffii*, collectively referred to as AMPS (the first letter for each species, capital letter when present, see also Table 2). Branches leading to *A. thaliana* receptors (in dark green) are labeled with the protein names. Other branches correspond to the basal flowering plant *Amborella trichopoda* (purple); the three monocots banana, rice, and maize (*Musa acuminata*, *Oryza sativa*, and *Zea mays*; yellow), and three dicots pink Shepherd's-purse, poplar, and tomato (*Capsella rubella*, *Populus trichocarpa*, and *Solanum lycopersicum*; cyan). LRR-RLKs with more than 20 LRRs from the recently sequenced algae *Penium margaritaceum* (the Penium Genome Database: http://bioinfo.bti.cornell. edu/cgi-bin/Penium/home.cgi; Jiao et al., 2020), *Spirogloea muscicola*, and *Mesotaenium endlicherianum* (Cheng et al., 2019) were used as outgroup (gray branches). Monophyletic groups of the same kind (i.e. either monocots, dicots, *A. trichopoda*, AMPS, or the outgroup) have been collapsed for clarity. Branches drawn with solid lines are well supported (SH-aLRT >0.95%), while dashed lines represent branches with lower support (see values for all branches in Supplemental File S1). KD: The lower part of the corresponding phylogenetic tree based on the KD illustrates ambiguities between the LRR and the KD trees and coincides with lower support (cf. eclipses; aLRT <0.85%). Branch lengths represent the average number of changes in amino acids per site. See Supplemental File S1 for the full KD tree and all branch values.

Table 3 Number of orthologues of 27 Arabidopsis LRR-RLKs subfamily XI in *A. trichopoda* and AMPS.

Vascular plants			Bryophytes			Common ancestor of vascular plants and bryophytes
Angiosperm		Lycophyte	Hornwort	Moss	Liverwort	
Arabidopsis thaliana [a]	*Amborella trichopoda*	***Selaginella moellendorffii***	*Anthoceros agrestis*	*Physcomitrium patens*	***Marchantia polymorpha***	
PEPR1	1	2[b]	2[b]	1[b]	0	PEPR
PEPR2						
GSO1	1				1	GSO
GSO2						
At1g17230	1	1	0	2	0	*At1g17230/ At2g33170/ At5g63930*
At2g33170	1					
At5g63930						
RGI5	1	3	0	0	2	RGI
RGI4	1					
RGI3						
RGI2	1					
RGI1						
HSL2	1	3	3	2	1	HSL
HSL1	1					
HAE						
HSL3	1					
CEPR1	1					
CEPR2	1					
RLK7	1					
IKU2						
TDR/PXY	1	2	1	0	1	PXY
PXL1	1					
PXL2						
BAM1	1	3	1	2	1	CLV1/BAM3
BAM2						
CLV1	1					
BAM3	1					
Total	18	14	7	7	6	6 or 7

[a]Bold represents novel genes in the eudicot lineage.
[b]Unsolved phylogeny

sequences are sister to the other angiosperm *PXL1/2* genes in the LRR tree but to the *PXY* genes in the KD tree, which perhaps reflects stronger purifying selection on the *PXL1/2* KD (Figure 3; Supplemental File S1 and Supplemental Table S2). While PXY remains as a single copy, the *PXL1/2* family expanded independently in eudicots and monocots, suggestive of their potential functional diversification in these lineages.

The *BAM/CLV1* gene family expanded markedly in the angiosperms with all the AMPS *BAM/CLV1* sequences in the sister grade. This phylogeny based on sequence information precludes a simple deduction of the ancestral state of vascular plant *BAM/CLV1* genes. Molecular genetic studies, however, started to shed light on the functional conservation between the vascular plant and bryophyte *BAM/CLV1* genes. Before further exploring the evolution of BAM/CLV1 and other LRXI receptors, we now turn to their partner, PTMPs to make a case for the receptor–ligand coevolution.

PTMPs in land plants

Members of 10 families of PTMPs with characteristic length and amino acid composition have been matched with LRR-RLKs (Table 1). Each family has been previously reviewed (Supplemental Table S2). Under the assumption that LRXI receptors are activated by PTMP ligands, conservation of LRXI orthologs in bryophytes and lycophytes triggers questions regarding the presence and the roles of signaling peptides in different lineages of land plants (Bowman et al., 2017; Liu et al., 2017). We, therefore, initiated our study by reviewing the PTMPs known to be present in nonflowering plants. PTMPs are generated from prepropeptides by posttranslational processing and amino acid modifications (Matsubayashi, 2014; Olsson et al., 2019). Gene and genome duplications have led to evolutionary changes, but the amino acid sequences of C-terminal bioactive peptides have been conserved. All known PTMPs were originally identified in angiosperms. Four families, all recognized by LRXI receptors, were reported to be conserved in nonseed plants: the H- and R-type of CLE peptides, each mainly recognized by TDR/PXY and BAM/CLV1; IDA interacting with HAE and HSL2 receptors; and ROOT GROWTH FACTOR (RGF)/GOLVEN (GLV)/CLE-LIKE (CLEL) peptides recognized by RGFR/RGI) receptors. Growing availability of nonseed plant genomes, however, demands reexamination of their phylogenetic distribution.

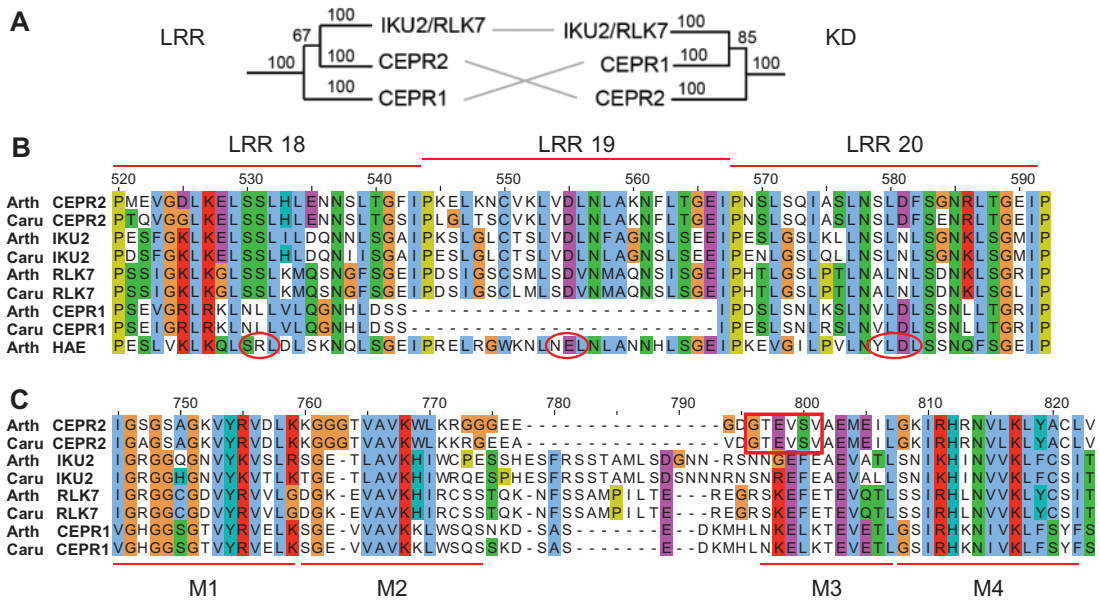

Figure 4 Differential sequence similarities in the ectodomains and the KDs of CEPR1, CEPR2, RLK7, and IKU2. A, Schematic representation of the phylogenetic difference between the LRR (left) and the KD (right) domain of CEPR1, CEPR2, RLK7, and IKU2 based on the phylogenies in Supplemental File S1. B, Alignment of three repeats of the ectodomain of the indicated receptors from *A. thaliana* and *Capsella rubella* corresponding to the Arabidopsis HAE repeats LRR18, LRR19, and LRR20. HAE residues interacting with the coreceptor SERK1 (Santiago et al. 2016) are encircled. C, Alignment of the N-terminal part the KD of the indicated receptors from *A. thaliana* and *C. rubella*. The first four conserved motives of the KD (Liu et al. 2016) are indicated below the alignment. The deviating M3 motif of CEPR2 is indicated by a red rectangle.

Furthermore, when we initiated our study, the remaining peptide families of CIFs, CEPs, PIP/PIP-LIKE, Peps, phytosulfokine [PSK], and PSY1) with members interacting with GSO, CEPR, RLK7, PEPR, PSKR, and PSYR receptors, respectively (for names see Table 1), had not (yet) been identified in nonseed plants. Hence, during this study, we have seen the need of searching for "missing" peptide ligand. However, this is not a trivial task since small open reading frames often are overlooked during the annotation of newly sequenced genomes, and prepropeptide sequences are highly variable except for the short region encoding the mature peptide.

It is generally difficult to find peptide homologs from distantly related species by Basic Local Alignment Search Tool for proteins (BLASTp; Altschul et al., 1990; https://blast.ncbi.nlm.nih.gov/Blast.cgi). Consequential false-negative problems may be circumvented by using exhaustive tBLASTn searches against whole-genome sequences or RNA collections translated in all six frames (Altschul et al., 1997). Limitation of the query to the most conserved part of the prepropeptide and limitation of BLASTp targets to coding sequences of less than 300 amino acids reduce the chance of getting false positive results. The search can be optimized for short query sequences, and the query further optimized using Pattern Hit Initiated (PHI) BLASTp and/or wildcards in positions where the amino acid residues are less conserved. Further options may be to test Machine Learning or Deep Learning methods that recently have been developed for peptide discovery (Plisson et al., 2020; Serrano, 2020; Zhang et al., 2020b).

The confidence of retrieved hits is usually low, and hits must therefore in all cases be analyzed for an appropriate prepropeptide length (<200 amino acids), the position of the assumed mature peptide sequence near or at the C-terminal end and the presence of a hydrophobic N-terminal secretion signal. When interpreting the sequences, it should be kept in mind that sequences from extant species may have diversified from ancestral sequences.

Using these strategies, we have identified putative peptide gene homologs in hornworts, lycophytes, ferns, and gymnosperms (Figure 1, indicated by dark triangles, see Supplemental Data Set S1). These data are presented below together with data on the presence of LRXI.

Coevolution of peptide ligands and their receptors

PEPR and GSO and the conservation of ligand-binding residues

PEPR and GSO were the receptors most closely related to the algal receptors used as outgroup in the phylogenetic analyses (Figure 3), although their functions in Arabidopsis may be angiosperm-specific. GSO is involved in the regulation of stem cell identity, control of seedling root growth, and establishment of the Casparian strip, which works as a diffusion barrier in the root vasculature (Tsuwamoto et al., 2008; Racolta, 2014; Doblas et al., 2017; Nakayama et al., 2017; Figure 3, A and B), while *A. thaliana* PEPR1 is involved in amplification of biotic and abiotic stress responses and

interact with endogenous stress-induced Pep peptides (Figure 1; Safaeizadeh and Boller, 2019). The phylogenetic distribution of the Pep ligands shown to interact with the AtPEPR is, however, limited to angiosperms, and even within flowering plants, Pep peptides show interfamily incompatibility. Peps of heterologous origin cannot be recognized as ligands due to the rapid coevolution of PEPR LRRs, and the *PEPR1–PEPR2* subclade has a higher ω value than most of the *LRXI* subclades (Table 2).

To elucidate why the PEPR ectodomain seems to be so closely related to the outgroup we used Repeat Conservation mapping which is a computational method that first removes the scaffold residues from the alignment of orthologous LRRs and thereafter calculates the conservation of the remaining residues (Koller and Bent, 2014). The result is presented as a heat map reflecting the degree of amino acid identity and similarity in a given position in a given repeat and can be interpreted with the help of 3D structure models. For PEPR, this analysis confirmed that the scaffold residues as well as putative coreceptor-interacting residues are conserved, but the majority of the Arabidopsis PEPR1 residues interacting with the peptide ligand Pep according to the solved crystal structure (Tang et al., 2015), are not conserved among PEPR orthologs from the species used in our studies (Figure 5, A, C, and E). This situation has likewise been found for another LRR-RLK receptor involved in defense, FLAGELLIN-SENSITIVE2 of subfamily XII, which, like PEPR, interacts with the coreceptor BAK1/SERK3 (Koller and Bent, 2014). This is in contrast to the conserved ligand-interacting residues of receptors involved in developmental processes, exemplified here by GSO1 interacting with the confirmed ligand CIF2 (Figure 5B, D and F; Okuda et al., 2020).

The solved crystal structures of LRXIs with ligands show that ligands interact along the inner surface of the ectodomain (Figure 5, A and B; Chakraborty et al., 2019). Hence, the length of the entire ectodomain must fit the length of the peptide ligand. PEPRs and GSOs illustrate this: they have 26 and 31 LRRs, respectively, while most LRXI has 21–23 repeats. Correspondingly, the peptides Pep and CIF interacting with PEPR and GSO are longer (17 and 20 amino acids, respectively) than the common 12–14 amino acids of the ligands of other LRXI receptors (Okuda et al., 2020; Tang et al., 2015).

RGIs and GSO and sulfated peptide ligands

In a search for ligands, it is important to take into account the number of LRRs and the length of the peptide ligand. Another factor to keep in mind when investigating potential peptide–receptor interactions is the presence of modified amino acid residues. Both CIFs and RGFs (Fernandez et al., 2013b) are known to have Asp (Asp, D) and a sulfated tyrosine (Tyr, Y) at the N-terminus (Table 1) and may be difficult to tell apart since both also are enriched in Pro residues, and often end with a C-terminal Asn (N). They do, however, differ in length (Figure 6, A and B; Table 1). Members of the RGF family serve diverse roles in root development, such as meristem homeostasis, gravitropism, and lateral root development (Fernandez et al., 2013b). The solved crystal structure of RGF1 and its receptor RGI3 (At4g26540; Song et al., 2016), also called RGFR3 (Shinohara et al., 2016), revealed that the DY^{sulf} residues are recognized by the amino acid motif RxGG of the receptor. Importantly, the RGI orthologs from different land plant lineages contain this motif, which distinguishes RGIs from other LRXI members (Supplemental Figure S1). RGFs have been suggested to be present in the lycophyte *S. moellendorffii* (Ghorbani et al., 2015), and signal through the RGI receptors (Shinohara et al., 2016; Song et al., 2016; Qian et al., 2018).

Whole-genome sequencing of nonflowering plants has facilitated the detection of GSO and RGI receptors (Bowman et al., 2017; Liu et al., 2017; Man et al., 2020), and we indeed retrieved ligand candidates from several species using Pattern-initiated BLASTp (PHI-BLASTp) with wildcards (Figures 1 and 6). The candidate CIF peptides show limited sequence identity (Figure 6A); however, this is also the case for the four Arabidopsis CIF peptides and the recently identified TWISTED SEED1 (TWS1) peptide which show differential affinities to GSO1 and GSO2 (Fiume et al., 2016; Barbosa et al., 2019; Doll et al., 2020; Okuda et al., 2020). TWS1 functions with the GSOs to form a functional cuticle around the developing embryo of Arabidopsis (Tsuwamoto et al., 2008).

To further clarify phylogenetic relationships among RGI homologs from different lineages, we bridged the evolutionary gap between Arabidopsis and *S. moellendorffii* by finding genes encoding highly similar receptors in the water ferns *Azolla filiculoide* and *Salvinia cucullata* and gymnosperm *Picea abies* (Norway spruce). In each of these species, two or three *RGI* genes were identified, one or two belonging to a clade that includes Arabidopsis *RGI5* and the other to a branch with three *A. trichopoda* and four Arabidopsis *RGIs* genes (Figure 6B), suggesting one gene duplication in the common ancestor of ferns and gymnosperms generated the RGI1–4 and RGI5 clades. In the RGI1–4 clade, at least three gene duplications are inferred: one in the seed plant lineage, generating two subgroups, both of which experienced duplication in the eudicot lineage. A previous study also indicated that gene copies increased through lineage-specific gene duplications (Man et al., 2020).

In Arabidopsis, Tyr sulfation regulates protein–protein interactions and affects receptor binding (Matsubayashi and Sakagami, 1996; Shinohara et al., 2016; Kaufmann and Sauter, 2019). The gene encoding the responsible enzyme, *TYROSYLPROTEIN SULFOTRANSFERASE* (*AtTPST*), was discovered in Arabidopsis (Komori et al., 2009). TPST-encoding genes are conserved across land plants and found in *S. moellendorffii*, *P. patens*, and *M. polymorpha*, as well in ferns and gymnosperms (Supplemental Figure S2) supporting the likely presence of sulfated signaling peptides and the plausible contribution of PTMs to the ligand diversity in early land plant evolution.

PSY1 and PSK are two sulfated peptides of 18 and 5 amino acids, respectively. PSY1 has an N-terminal DY^{sulf}, and promotes cellular proliferation and expansion (Amano et al.,

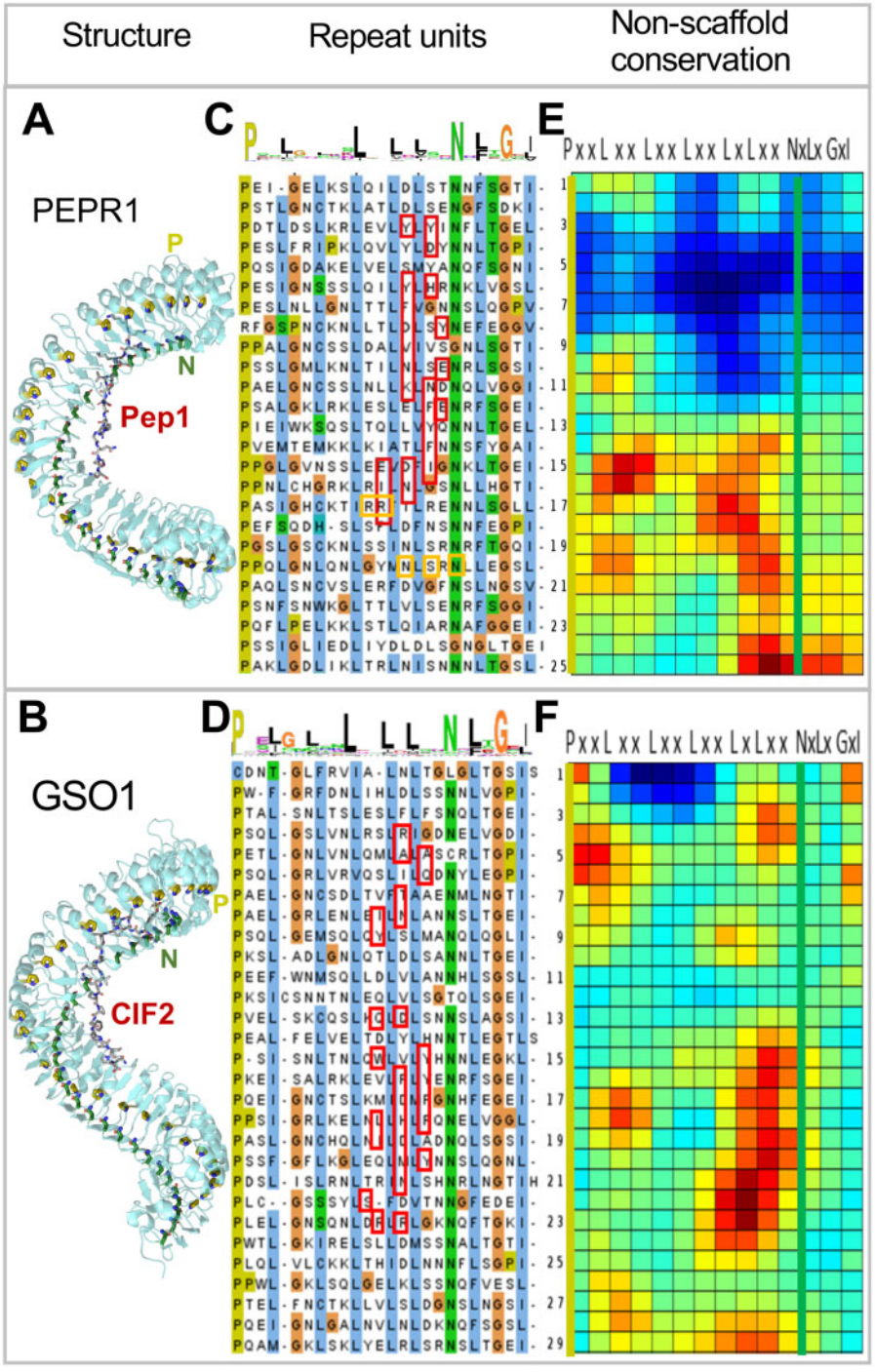

Figure 5 Conservation of LRR scaffold residues and ligand-interacting amino acids. A and B, Crystal structure of PEPR1 (PDBid:5GR8) with the ligand Pep1 (Tang et al., 2015) and GSO1 (PDBid: 6S6Q) with the ligand CIF2 (Okuda et al., 2020), respectively, interacting along the inner side of the LRR structures. The conserved Pro (P) and Asn (N) residues of the scaffolds are highlighted in mustard and green colors. C and D, Alignment of the 24 amino acids long leucine-rich repeat units of PEPR1 and GSO1 with coloration based on amino acid properties according to the default Clustal X color scheme. Above a WebLogo consensus sequence (http://weblogo.berkeley.edu/) visualizing the conservation of the scaffold residues. Residues of the LRR that according to the crystal structures shown in (A) and (D) interact with residues of the respective peptides, are marked with red rectangle, and amino acids of PEPR1 interacting with the co-receptor BAK1/SERK3 are marked with orange squares. E and F, Heat maps generated using Repeat Conservation Mapping (http://www.bentlab.russell.wisc.edu/main/main.php) reflecting the degree of amino acid identify and similarity in a given position (*X*-axis) and a given repeat (*Y*-axis) for the nonscaffold residues of PEPR and GSO orthologs, respectively. The most conserved residues are in red and least conserved in blue. The position of the conserved Ps and Ns of the scaffolds are indicated by mustard and green colored vertical lines, respectively. Note that the conserved amino acid residues of PEPR orthologs (LRR 13 and onward) are not coinciding with the majority of the amino acids of PEPR1 interacting with the Pep1 ligand (LRRs 2–13), but rather with the coreceptor-interacting residues. In contrast, ligand-interacting residues overlap with a substantial number of conserved residues in GSO orthologs (LRRs 14–23).

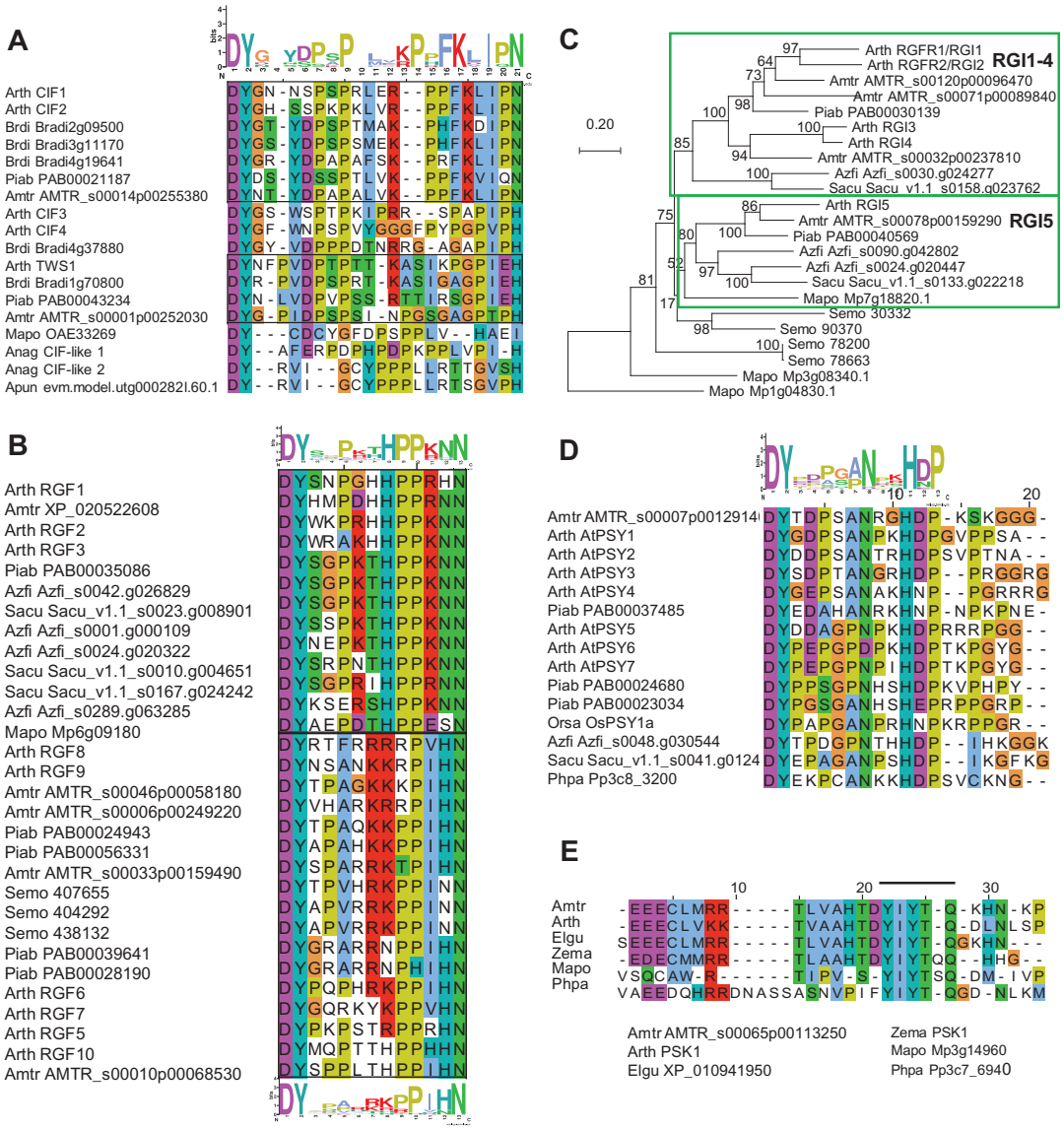

Figure 6 Conservation of sulfated peptides. A, Alignment of the CIF and TWS1-related peptides. Above the alignment shown is the sequence logo of the top seven sequences (boxed), which represent canonical CIFs. TWS1 and its putative orthologs are also boxed. B, Alignment of the presumptive RGF peptides, with amino acids colored according to chemical properties. RGFs are classified into two distinct groups (boxed), represented by the sequence logos above and below the alignment. C, Phylogenetic tree based on the ectodomain amino acid sequences of RGIs in Arabidopsis, and their orthologues in *A. trichopoda*, the gymnosperm *P. abies*, the ferns *A. filiculoides* and *S. cucullata*, the lycophyte *S. moellendorffii* and the bryophyte *M. polymorpha* constructed by the maximum likelihood method. The GSO homolog of *M. polymorpha* (Mp1g04830.1) from the neighboring clade was used as outgroup. Bootstrap values were obtained from 500 bootstrap replicates and indicated at the nodes as percentages. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. D, Alignment of the presumptive PSY peptides, with a sequence logo of the first 13 amino acids above the alignment. E, Alignment of the PSK1 mature peptide (YIYTQ) and adjacent conserved motif. Sequence logos were generated using WebLogo 2.8.2 (https://weblogo.berkeley.edu). Conserved amino acids are colored based on amino acid properties according to the default Clustal X color scheme. Amtr, *Amborella trichopoda*; Anag, *Anthoceros agrestis*; Anpu, *Anthoceros punctatus*; Arth, *Arabidopsis thaliana*; Azfi, *Azolla filiculoides*; Brdi, *Brachypodium distachyon*; Elgu, *Elaeis guineensis*, Mapo, *Marchantia polymorpha*; Orsa, *Oryza sativa*; Phpa, *Physcomitrium patens*; Piab, *Picea abies*; Sacu, *Salvinia cucullata*; Semo, *Selaginella moellendorffii*, Zema, *Zea maize*.

2007). The mature PSK peptide is only five amino acids long (YIYIQ) and promotes cell growth, acts in the root apical meristem, contributes to pollen tube guidance, and integrates growth and defense (Sauter, 2015; Kaufmann and Sauter, 2019). We identified putative orthologs both for PSY1 and PSK in *P. patens*, *M. polymorpha* as well as *P. abies* (Figure 6, D and E). *PSK*-like genes and bioactivities have previously been detected in gymnosperms (Igasaki et al., 2003; Wu et al., 2019), but surprisingly, we did not find any ortholog of the Arabidopsis PSY1-receptor in the bryophytes. In contrast to all the subfamily XI receptors discussed here, the two PSK-RECEPTORs (PSKRs) of Arabidopsis belong to the subfamily X, of which binding to small ligands is facilitated by a so-called island domain in the LRR structure that does not

fit in with the rest of the repeats. In the PSK-like sequences found in *P. patens* and *M. polymorpha*, the Asp (D) residue preceding the YIYIQ sequence is missing, but a conserved N-terminal region can still be recognized (Figure 6E). Such conserved sequences outside the assumed mature peptide might be involved in precursor processing and peptide maturation steps, which have been less characterized in plant peptidic signals.

The HSL receptors and their ligands evolved through successive changes

The angiosperm HAE-HSL receptors showcase dynamic processes of their coevolution with peptide ligands. In Arabidopsis the receptors HAE and HAE-LIKE2 (HSL2) control cell separation events through the IDA ligand which is a 12 amino acid proline-rich peptide with a concerved histidine–asparagine (His–Asn, HN) or HH at the C-terminus, residues that have been shown to interact with two closely positioned arginine (Arg, R) residues found in many LRXI receptors (Table 1; Hou et al., 2014; Vie et al., 2015; Stø et al., 2015; Song et al., 2016; Shi et al., 2019). As stated previously, the four AMPS have LRXI in the HSL clade. This and the recent identification of orthologs of IDA in *M. polymorpha* (Bowman et al., 2017), encouraged us to search for IDA in the other nonflowering plants. We identified potential IDA orthologs in *S. moellendorffii* and *Anthoceros* species, however, not in *P. patens* (Figures 1 and 7).

To substantiate that the MpIDA and SmIDA peptides could be the ligands of MpHSL2 and SmHSL2, respectively, we took advantage of the solved crystal structure of synthetic AtIDA peptide bound to its receptor AtHAE (Santiago et al., 2016). Using this structure, we have recently successfully generated a 3D model of IDA-HSL2 orthologs from oil palm (Shi et al., 2019)). Using the same modeling strategy for *M. polymorpha* and *S. moellendorffii*, we demonstrate that a substantial fraction of the AtHAE amino acids generating hydrogen bonds with IDA residues, are identical in MpHSL2 and SmHSL2 (Figure 7, A and B; Supplemental Figure S3). The potential overall hydrogen bonding interactions between the modeled MpHSL2 and SmHSL2 receptors and their corresponding IDA peptides (MpIDA1-2 and SmIDA1-3) support receptor–peptide binding (Figure 7, B and C). Hydroxylation of the centrically positioned Pro residues has been shown to increase the binding of IDA and HAE/HSL2 in *A. thaliana* (Santiago et al., 2016), and accordingly, hydroxylation of this P in the in silico model generated more hydrogen bonds between the receptor and the IDA ligand (Figure 7, C and D). Crucial residues, in particular, the central Pro and the C-terminal His–Asn, are also conserved in the amino acid sequence of the MpIDA1-2 and SmIDA peptides (Figure 7B). Thus, the generated models are consistent with a function as ligand–receptor pairs in early land plant lineages. It remains to be demonstrated whether such a pair plays a role in cell wall remodeling and/or cell separation in nonflowering plants.

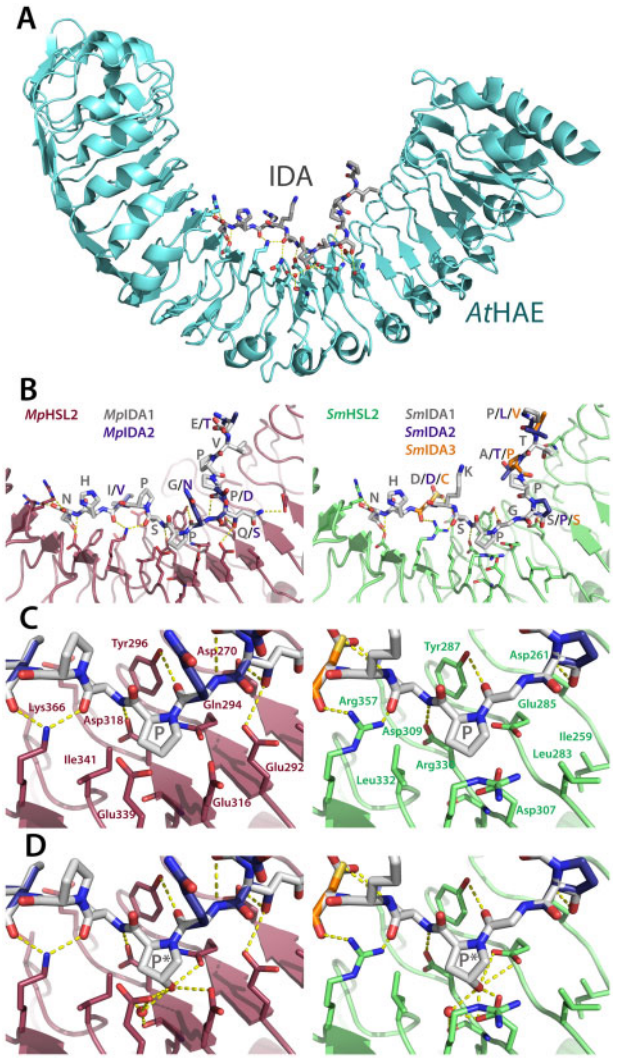

Figure 7 Modeling of the interaction between putative IDA ligands and HSL2 receptors of *M. polymorpha* and *S. moellendorffii*. A, The AtHAE-AtIDA crystal structure (PDBid:5ixq; Santiago et al. 2016) with the IDA peptide lined along the inner face of the LRR structure. B, Overall models of the interaction between *M. polymorpha* and *S. moellendorffii* putative HSL2 receptors (MpHSL2 in magenta and SmHSL2 in green) with, respectively, two and three *M. polymorpha* and *S. moellendorffii* superimposed putative IDA peptides (gray backbones) built on the AtHAE-AtIDA structure using SWISS-MODEL (Arnold et al., 2006). Note in particular receptor interaction with the Asn (N) at the C-terminal end of the peptides. C, Close-up view of the central parts of the respective receptor models and the surrounding hydrogen bonding network, with a central Pro (P) in the ligands. D, Close-up view as in (C), however, with hydroxylation of the central Pro (P*), which facilitates formation of additional hydrogen bonds. Central amino acids of the receptors, as well as the peptides, are shown as sticks and colored by atom type. Water molecules are shown as red spheres, and modeled based on coordinates from the AtHAE-AtIDA crystal structure. Hydrogen bonds are depicted as dotted lines (yellow). Residues involved in hydrogen bonding to the peptides are depicted with three-letter symbols in colors according to the respective structures. The peptide residues are shown in one-letter symbols. All structure figures were prepared using PyMOL (Schrödinger, LLC).

Interestingly, other receptors in the HSL clade, namely, CEPRs and RLK7, do not bind IDA peptides, but instead have the closely related CEPs and PIP–PIPLs as their ligands. They are similar to IDA peptides with respect to the size and amino acid composition, possibly reflecting structural constraints to stabilize homologous receptor–ligand interactions. A notable difference is that most CEPs and PIP–PIPLs lack the C-terminal HN or HH. Instead, CEPs and PIP–PIPLs share the C-terminal GxGH motif (Vie et al., 2015; Table 1). Such deviation could be explained by recent origins of both receptors and ligands. The origin of CEPs in gymnosperms (Delay et al., 2013; Roberts et al., 2013) predates the emergence of PIP–PIPLs in angiosperms. Elucidating whether gymnosperms have CEPR orthologs and deorphanizing HSL3 and IKU2 would help us trace coevolutionary processes of receptor homologs and their potentially homologous ligands. Considering what we now know about the *HSL* clade, gene duplications seem to have resulted in novel functions.

CLE peptides and evolution of peptide signaling

The CLE family peptides illustrate contrasting themes. The 12–13 amino acid CLE peptides are proline (Pro, P)-rich and characteristically have the above-mentioned C-terminal HN or HH (Table 1). CLE peptides are classified into two groups by the N-terminal residue: the R (arginine, Arg) and H groups. Many R-CLEs interact with CLAVATA1 (CLV1) or BARELY ANY MERISTEM (BAM) receptors involved in meristem maintenance in Arabidopsis (DeYoung and Clark, 2008; Ogawa et al., 2008), while the three H-CLE peptides interact with PXY/TDR involved in vascular differentiation in Arabidopsis (Hirakawa et al., 2010). The CLE peptides are the first PTMPs from bryophytes that have been assigned with biological functions. These studies revealed that the CLE peptide signaling has evolutionarily conserved roles from *P. patens* and *M. polymorpha* to angiosperms (Hata and Kyozuka, 2021). One of the known key roles of the CLE signaling in angiosperms is to regulate cell proliferation in the sporophytic (2*n*) meristems, which grow indeterminately. In the bryophytes, the sporophyte shows determinate growth while indeterminate meristems are present in the gametophyte body (1*n*). In *M. polymorpha* an R-type peptide (MpCLE2) functions in a paracrine manner as a haploid stem cell-promoting signal in dichotomous branching of the thallus and is genetically dependent on the MpCLV receptor (Hirakawa et al., 2020). In *P. patens*, the R-type PpCLE peptide with its PpCLV receptor is involved 3D growth by determination of the orientation of stem cell division plane (Whitewoods et al., 2018).

To our surprise, several Arabidopsis CLE peptides have been found to interact with LRRs of different clades. The AtCLE9/CLE10 peptides interact with both the HSL1 and BAM1 receptors, depending on tissue type (Qian et al., 2018). AtCLE14 peptide signals through AtPEPR2 (Gutierrez-Alanis et al., 2017). While PEPs and CLEs differ in length, the C-terminal 12 amino acids of mature PEPs share sequence similarity to CLEs (Table 1), offering a possible explanation for the recognition mechanism. Another CLE peptide, AtCLE45, which differs from other CLE peptides by a number of arginine (Arg, R) residues near its N-terminus, is perceived by canonical CLE receptors as well as an RGF receptor homolog, AtRGI4/SKM2 (Endo et al., 2013; Hazak et al., 2017).

An MpTDR receptor, is on the other hand confined to the proliferative activity of gametophytic meristem and affects the overall size of reproductive organs (gametangiophore; Hirakawa et al., 2019). Functions of the H- and R-type CLE peptides in *Anthoceros* remain to be revealed (Li et al., 2020; Zhang et al., 2020a; Figure 1). The studies in the moss and liverwort support the idea that peptide–receptor modules are conserved and employed in a parallel manner, although the evolutionary relationship between the bryophyte gametophytic meristems and the vascular plant sporophytic meristems is still debated (Bowman, 2013). A key question is whether conserved mechanisms at cellular levels (e.g. downstream signaling events and target pathways) underpin functional conservation between analogous systems of distantly related species.

Toward decoding peptide signals

The "unconventional" ligand–receptor pairs discussed above raise the question of how the specificity of ligand–receptor pairs evolves. It is of note that genes encoding CLE peptides highly similar to AtCLE9/CLE10 are present in both *M. polymorpha* and *P. patens* (PpCLE5/CLE6; Whitewoods et al., 2018; Figure 8). Our phylogenetic investigations suggest that founder receptors, representing the seven clades, by the onset of the angiosperm era had increased about three times, and in each of the clades, a new gene has evolved during dicot evolution. We therefore speculate that during the early evolution of peptide signaling, there were fewer peptides, and receptors were less specific. Accordingly, large-scale clustering analyses of land plant CLE peptides based on their entire prepropeptide sequences found a smaller number of clusters in AMPS, indicative of the diversification of CLE peptide sequences, which could have resulted in changes in the specificity of ligand–receptor interactions during land plant evolution (Goad et al., 2017). Hirakawa et al. (2017) found that a synthetic CLE peptide, KIN named after the K (2nd), I (10th), and N (12th) residues crucial for activity, exerts both R-type and H-type CLE activities and interact directly with both the CLV1 and the TDR receptors (Figure 8; Hirakawa et al., 2017), illustrating the potential for a broader specificity in ligand–receptor interactions. It is of critical importance to reveal what underlies common features and family-specific characteristics in order to understand both causes (e.g. homology or convergence) and consequences (e.g. family expansion, birth of new ligands, or signaling crosstalk) of evolutionary changes in signaling peptides.

Perspectives

Genetics and nonflowering model plants

In the 20 years since Shiu and Bleecker categorized the LRR-RLKs of Arabidopsis (Shiu and Bleecker, 2001), our

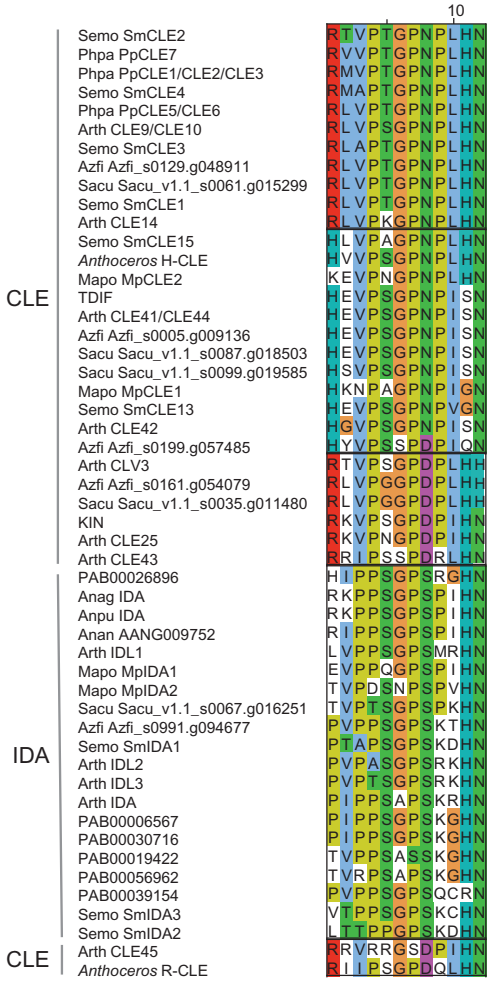

Figure 8 Conservation of CLE and IDA peptides. Alignment of mature CLE and IDA peptides in land plants and a synthetic peptide, KIN. Note shared core residues (Pro-Ser-Gly-Pro (PSGP) and C-terminal end (HN) between IDA and some of the CLE peptides, and the presence of almost identical R-CLE, H-CLE, and IDA peptides in AMPS, ferns, and Arabidopsis. Conserved amino acids are colored according to chemical properties using the default Clustal X color scheme. Amtr, *Amborella trichopoda*; Anag, *Anthoceros agrestis*; Anpu, *Anthoceros punctatus*; Arth, *Arabidopsis thaliana*; Azfi, *Azolla filiculoides*; Brdi, *Brachypodium distachyon*; Mapo, *Marchantia polymorpha*; Orsa, *Oryza sativa*; Phpa, *Physcomitrium patens*; PAB, *Picea abies*; Sacu, *Salvinia cucullata*; Semo, *Selaginella moellendorffii*.

knowledge of their functions and their confirmed ligands has gradually increased. Principally it is expected that disruption of the genes encoding a receptor or its ligand display the same mutant phenotype. The CLAVATA1 (CLV1) receptor and the CLV3 peptide ligand were discovered at the end of the last century by their spectacular meristem phenotypes (Clark et al., 1995, 1997). A double mutant of the closely related receptors HAE and HSL2 was needed to disclose their involvement in floral organ abscission, like their peptide ligand IDA (Butenko et al., 2003; Cho et al., 2008; Stenvik et al., 2008b). Furthermore, triple to quintuple mutants were necessary to disclose the involvement of the RGI receptors in root development, with RGFs peptides as ligands (Shinohara et al., 2016; Song et al., 2016). Thus, functional redundancy on the receptor side as well as the peptide side has made it complicated to identify ligand–receptor pairs through genetics. However, in nonseed plants that often have fewer duplicated genes, for example, *M. polymorpha*, and with the establishment of CRISPR-Cas mutagenesis, a genetic approach for the identification of peptide ligand receptor-pairs and their biochemical and biological functions, is becoming feasible. There is thus a need for developing more nonseed, genetically tractable model plants (Rensing, 2017). For instance, ligand(s) or function(s) have not been reported for the LRXI clade including the three Arabidopsis *Atan genes* (Figure 3C); studying its moss orthologs could provide first functional insight into this clade.

Families and orphans

A number of receptors and members of peptide families are orphans, in the sense that their signaling partners are unknown. This may result from the accumulation of mutations in duplicated genes that weakened ligand–receptor interaction. An alternative (and likely) explanation is that more specific genetic, molecular, and biochemical factors and conditions for peptide-mediated cell-to-cell communication remain to be discovered. In line with this suggestion, CLE9/10 of Arabidopsis bind HSL1 efficiently in the presence of the small multifaceted coreceptors SERK1, -2 or -3, but prefer BAM1 in the absence of SERK (Qian et al., 2018). To dig deeper into the evolution of peptide signaling, we recommend matching phylogenetic, functional, and structural data to reveal additional interactions with other XI receptors (Smakowska-Luzan et al., 2018) and coreceptors as key partners in receptor complexes in diverse peptide signaling pathways (Ma et al., 2016; Hohmann et al., 2017; He et al., 2018; Liang and Zhou, 2018).

Considering that a fourth of the present-day XI receptors of Arabidopsis, with seemingly angiosperm-specific functions, has a history all the way back to the ancestor of the bryophytes and vascular plants, expanding our knowledge of nonflowering plants will give a new perspective on the roles and evolution of cell-to-cell signaling. In addition to CLE peptide ligands and their receptors (BAM/CLV and TDR; Whitewoods et al., 2018; Hirakawa et al., 2019; Hirakawa et al., 2020), our approach has disclosed the early occurrence of orthologs of GSO and RGI receptors and their CIF and RGF ligands. In addition, direct interactions between HSL2 and IDA orthologs of *M. polymorpha* and *S. moellendorffii* have been substantiated, taking advantage of the solved 3D structure of the Arabidopsis peptide–receptor pair (Figure 7). We have also bridged the gap between AMPS and Arabidopsis by identifying likely orthologs of *LRXI* in ferns and gymnosperms (Figure 6C). These may be starting points for investigating the roles of peptide signaling and LRXI evolution prior to the emergence of angiosperms. An approach similar to what we have shown here with the integration of phylogenetic, functional, and structural information, should facilitate the disentanglement of distinct evolutionary

histories of peptide ligands and their receptors. One study along these lines has recently been published on the BAM/CLV1/TDR receptors and CLE peptides (Cammarata and Scanlon, 2020).

Two ancient types of signaling PTMPs

Interestingly, the GSO/RGI clades and the HSL/TDR/BAM/CLV clades seem to have ligands representing two different groups—those with sulfated Tyr and the CLE/IDA-like peptides, respectively. Founder genes for both were possibly present in the most recent common ancestor of vascular plants and bryophytes. Alternatively, convergent evolution cannot be dismissed and might account for the highly variable central part of peptide precursor sequences. Signaling peptides can evolve not only from changes in the primary sequences, but also through changes in proteolytic processing of prepropeptides, or by PTM enzymes conferring Tyr sulfation or Pro hydroxylation. PTM processes tend to be irreversible and hence require precise regulation, pointing to the possible contribution of PTMs in the diversification of peptide ligands and in defining the specificity of ligand–receptor interaction since the early stages of land plant evolution. Our structure-based modeling of IDA-HSL2 is consistent with an ancient origin of peptide hydroxylation. Structural modeling and analyses of bioactive forms of PTMPs in AMPS will be critical for confirming predictions of molecular interactions.

New functions and new perspectives

Land plant evolution was influenced by a broadening spectrum of receptors as more complex organs evolved and interspecies interactions increased. However, except for the bryophyte CLE signaling pathways, biological roles have only been assigned for peptide LRXI signaling in angiosperms, and most often have been studied in angiosperm-specific organs. Where and when a signaling system triggers a given outcome (e.g. cell division, cuticle formation, or cell separation) may be the crux of the matter.

Not all the known signaling peptide families are conserved in nonseed plants, including bryophytes (Table 2). We know very little to nothing about peptide signaling pathways that uniquely evolved in these lineages. As such, much remains to be discovered to truly appreciate the diversity of land plant peptide signaling and to postulate testable hypotheses regarding its origin and evolution. In this Perspective, using diverse in silico methods, we broadened our knowledge on the molecular inventory of peptide-signaling components and presented hypotheses to be addressed. A very important next step in tracing the evolutionary origin and history of land plant peptide-signaling pathways will be to investigate thoroughly the temporal and spatial expression patterns of LRXI receptors and their putative peptide ligands, and to study downstream molecular events in diverse species.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Conservation of ligand-binding residues in RGI receptors. Support for Figure 6.

Supplemental Figure S2. Phylogenetic analysis of TYROSYLPROTEIN SULFOTRANSFERASE (TPST). Support for Figure 6.

Supplemental Figure S3. Conservation of amino acids residues involved in ligand binding. Support for Figure 7.

Supplemental Table S1. Major cysteine-rich peptide families in land plants. Support for Table 1.

Supplemental Table S2. References for the signaling peptides. Support for Table 1.

Supplemental File S1. Phylogeny for subfamily XI LRR-RLKs based on the KD and the LRR ectodomain. Support for Figure 2.

Supplemental File S2. Alignment of LRRs for phylogenetic analyses.

Supplemental File S3. Alignment of KDs for phylogenetic analyses.

Supplemental File S4. Phylogenetic tree file for LRRs.

Supplemental File S5. Phylogenetic tree file for KDs.

Supplemental Data Set 1. Names and IDs for proteins and genes presented in Figure 2 and Tables 1 and 2. Excel format.

Acknowledgments

We apologize to authors of literature not discussed due to space constraints. We thank three anonymous reviewers for helpful comments that improved the manuscript.

Funding

Investigations by C.F. and S.S. were supported by Japan Society for the Promotion of Science London (KAKENHI 18H05487, 20H00422, and 20KK0135 to S.S. and 20K06770 to C.F.) and research grants from the Nakatsuji Foresight Foundation and the Sumitomo Foundation to C.F. The Research Council off Norway has supported RBA (grant RCN312785).

Conflict of interest statement. The authors declare no conflict of interest.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215**: 403–410

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25

Amano Y, Tsubouchi H, Shinohara H, Ogawa M, Matsubayashi Y (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in Arabidopsis. Proc Natl Acad Sci U S A **104**: 18333–18338

Amborella Genome Project (2013) The Amborella genome and the evolution of flowering plants. Science **342**: 1241089

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics (Oxford, England) **22**: 195–201

Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, dePamphilis C, Albert VA, Aono N, Aoyama T, Ambrose BA, et al. (2011) The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science **332**: 960–963

Barbosa ICR, Rojas-Murcia N, Geldner N (2019) The Casparian strip—one ring to bring cell biology to lignification? Curr Opin Biotechnol **56**: 121–129

Bowman JL (2013) Walkabout on the long branches of plant evolution. Curr Opin Plant Biol **16**: 70–77

Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, Yamaoka S, Nishihama R, Nakamura Y, Berger F, et al. (2017) Insights into land plant evolution garnered from the marchantia polymorpha genome. Cell **171**: 287–304.e215

Butenko MA, Patterson SE, Grini PE, Stenvik GE, Amundsen SS, Mandal A, Aalen RB (2003) *INFLORESCENCE DEFICIENT IN ABSCISSION* controls floral organ abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants. Plant Cell **15**: 2296–2307

Cammarata J, Scanlon MJ (2020) A functionally informed evolutionary framework for the study of LRR-RLKs during stem cell maintenance. J Plant Res **133**: 331–342

Chakraborty S, Nguyen B, Wasti SD, Xu G (2019) Plant leucine-rich repeat receptor kinase (LRR-RK): structure, ligand perception, and activation mechanism. Molecules **24**: 3081

Chen T, Wang B, Wang F, Niu G, Zhang S, Li J, Hong Z (2020) The evolutionarily conserved serine residues in BRI1 LRR motifs are critical for protein secretion. Front Plant Sci **11**: 32

Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, Sun W, Li X, Xu Y, Zhang Y, et al. (2019) Genomes of subaerial zygnematophyceae provide insights into land plant evolution. Cell **179**: 1057–1067.e1014

Cho SK, Larue CT, Chevalier D, Wang H, Jinn TL, Zhang S, Walker JC (2008) Regulation of floral organ abscission in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A **105**: 15629–15634

Clark SE, Running MP, Meyerowitz EM (1995) CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. Development **121**: 2057–2067

Clark SE, Williams RW, Meyerowitz EM (1997) The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. Cell **89:** 575–585

de Vries J, Curtis BA, Gould SB, Archibald JM (2018) Embryophyte stress signaling evolved in the algal progenitors of land plants. Proc Natl Acad Sci U S A **115**: E3471–E3480

Delay C, Imin N, Djordjevic MA (2013) CEP genes regulate root and shoot development in response to environmental cues and are specific to seed plants. J Exp Bot **64**: 5383–5394

Delwiche CF, Cooper ED (2015) The evolutionary origin of a terrestrial flora. Curr Biol **25**: R899–910

DeYoung BJ, Clark SE (2008) BAM receptors regulate stem cell specification and organ development through complex interactions with CLAVATA signaling. Genetics **180**: 895–904

Doblas VG, Smakowska-Luzan E, Fujita S, Alassimone J, Barberon M, Madalinski M, Belkhadir Y, Geldner N (2017) Root diffusion barrier control by a vasculature-derived peptide binding to the SGN3 receptor. Science **355**: 280–284

Doll NM, Royek S, Fujita S, Okuda S, Chamot S, Stintzi A, Widiez T, Hothorn M, Schaller A, Geldner N, et al. (2020) A two-way molecular dialogue between embryo and endosperm is required for seed development. Science **367**: 431–435

Edgar Robert C (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research **32**: 1792–1797

Endo S, Shinohara H, Matsubayashi Y, Fukuda H (2013) A novel pollen-pistil interaction conferring high-temperature tolerance during reproduction via CLE45 signaling. Curr Biol **23**: 1670–1676

Fernandez A, Drozdzecki A, Hoogewijs K, Nguyen A, Beeckman T, Madder A, Hilson P (2013a) Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. Plant Physiol **161**: 954–970

Fernandez A, Hilson P, Beeckman T (2013b) GOLVEN peptides as important regulatory signaling molecules of plant development. J Exp Bot **64**: 5263–5268

Fiume E, Guyon V, Remoué C, Magnani E, Miquel M, Grain D, Lepiniec L (2016) TWS1, a novel small protein, regulates various aspects of seed and plant development. Plant Physiol **172**: 1732

Frangedakis E, Shimamura M, Villarreal J, Li F, Tomaselli M, Waller M, Sakakibara K, Renzaglia Karen S, Szövényi P (2021) The hornworts: morphology, evolution and development. The New Phytologist. **229**: 735–754

Friedman WE, Williams JH (2004) Developmental evolution of the sexual process in ancient flowering plant lineages. Plant Cell **16 (Suppl)**: S119–S132

Gensel PG (2008) The earliest land plants. Annu Rev Ecol Evol Systemat **39**: 459–477

Ghorbani S, Lin YC, Parizot B, Fernandez A, Njo MF, Van de Peer Y, Beeckman T, Hilson P (2015) Expanding the repertoire of secretory peptides controlling root development with comparative genome analysis and functional assays. J Exp Bot **66**: 5257–5269

Goad DM, Zhu C, Kellogg EA (2017) Comprehensive identification and clustering of CLV3/ESR-related (CLE) genes in plants finds groups with potentially shared function. New Phytol **216**: 605–616

Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK (2002) ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. Cell **111**: 803–814

Grienenberger E, Fletcher JC (2015) Polypeptide signaling molecules in plant development. Curr Opin Plant Biol **23**: 8–14

Gutierrez-Alanis D, Yong-Villalobos L, Jimenez-Sandoval P, Alatorre-Cobos F, Oropeza-Aburto A, Mora-Macias J, Sanchez-Rodriguez F, Cruz-Ramirez A, Herrera-Estrella L (2017) Phosphate starvation-dependent iron mobilization induces CLE14 expression to trigger root meristem differentiation through CLV2/PEPR2 signaling. Dev Cell **41**: 555–570 e553

Hata Y, Kyozuka J (2021) Fundamental mechanisms of the stem cell regulation in land plants: lesson from shoot apical cells in bryophytes. Plant Mol Biol, 10.1007/s11103-021-01126-y

Hazak O, Brandt B, Cattaneo P, Santiago J, Rodriguez-Villalon A, Hothorn M, Hardtke CS (2017) Perception of root-active CLE peptides requires CORYNE function in the phloem vasculature. EMBO Rep **18**: 1367–1381

He Y, Zhou J, Shan L, Meng X (2018) Plant cell surface receptor-mediated signaling – a common theme amid diversity. Journal of Cell Science **131** 10.1242/jcs.209353

Hetherington AJ, Dolan L (2019) Evolution: diversification of angiosperm rooting systems in the early cretaceous. Curr Biol **29**: R1081–R1083

Hirakawa Y, Fujimoto T, Ishida S, Uchida N, Sawa S, Kiyosue T, Ishizaki K, Nishihama R, Kohchi T, Bowman JL (2020) Induction of multichotomous branching by CLAVATA peptide in Marchantia polymorpha. Curr Biol **30**: 3833–3840.e4

Hirakawa Y, Kondo Y, Fukuda H (2010) Regulation of vascular development by CLE peptide-receptor systems. J Integr Plant Biol **52**: 8–16

Hirakawa Y, Torii KU, Uchida N (2017) Mechanisms and strategies shaping plant peptide hormones. Plant Cell Physiol **58**: 1313–1318

Hirakawa Y, Uchida N, Yamaguchi YL, Tabata R, Ishida S, Ishizaki K, Nishihama R, Kohchi T, Sawa S, Bowman JL (2019) Control of proliferation in the haploid meristem by CLE peptide signaling in Marchantia polymorpha. PLoS Genet **15**: e1007997

Hohmann U, Lau K, Hothorn M (2017) The structural basis of ligand perception and signal activation by receptor kinases. Annual Review of Plant Biology **68**: 109–137

Hosseini S, Schmidt EDL., Bakker FT (2020) Leucine-rich repeat receptor-like kinase II phylogenetics reveals five main clades throughout the plant kingdom. Plant J **103**: 547–560

Hou S, Wang X, Chen D, Yang X, Wang M, Turra D, Di Pietro A, Zhang W (2014) The secreted peptide PIP1 amplifies immunity through receptor-like kinase 7. PLoS Pathog **10**: e1004331

Igasaki T, Akashi N, Ujino-Ihara T, Matsubayashi Y, Sakagami Y, Shinohara K (2003) Phytosulfokine stimulates somatic embryogenesis in Cryptomeria japonica. Plant Cell Physiol **44**: 1412–1416

Ishizaki K (2017) Evolution of land plants: insights from molecular studies on basal lineages. Biosci Biotechnol Biochem **81**: 73–80

Jiao C, Sørensen I, Sun X, Sun H, Behar H, Alseekh S, Philippe G, Palacio Lopez K, Sun L, Reed R, et al. (2020) The Penium margaritaceum genome: hallmarks of the origins of land plants. Cell **181**: 1097–1111.e1012

Jill Harrison C (2017) Development and genetics in the evolution of land plant body plans. Philos Trans R Soc Lond B Biol Sci **372**: 20150490

Jones P (2014) InterProScan 5: genome-scale protein function classification Bioinformatics **30**: 1236–1240

Kaufmann C, Sauter M (2019) Sulfated plant peptide hormones. J Exp Bot **70**: 4267–4277

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature **389**: 33–39

Koller T, Bent AF (2014) FLS2-BAK1 extracellular domain interaction sites required for defense signaling activation. PloS One **9**: e111185–e111185

Komori R, Amano Y, Ogawa-Ohnishi M, Matsubayashi Y (2009) Identification of tyrosylprotein sulfotransferase in Arabidopsis. Proc Natl Acad Sci U S A **106**: 15067–15072

Lamesch P, Berardini Tanya Z, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander Debbie L, Garcia-Hernandez M, et al. (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research **40**: D1202–10

Lang D, Ullrich Kristian K, Murat F, Fuchs J, Jenkins J, Haas Fabian B, Piednoel M, Gundlach H, Van Bel M, Meyberg R, et al. (2018) The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. The Plant Journal : For Cell and Molecular Biology **93**: 515–533

Lehti-Shiu MD, Zou C, Hanada K, Shiu S-H (2009) Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. Plant Physiol **150**: 12–26

Li F-W., Brouwer P, Carretero-Paulet L, Cheng S, de Vries J, Delaux P-M., Eily A, Koppers N, Kuo L-Y., Li Z, et al. (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants **4**: 460–472

Li FW, Nishiyama T, Waller M, Frangedakis E, Keller J, Li Z, Fernandez-Pozo N, Barker MS, Bennett T, Blázquez MA, et al. (2020) Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants **6**: 259–272

Liang X, Zhou J (2018) Receptor-like cytoplasmic kinases: Central players in plant receptor kinase-mediated signaling. Annual Review of Plant Biology **69**: 267–299. 10.1146/annurev-arplant-042817-040540. 29719165.

Ligrone R, Duckett JG, Renzaglia KS (2012) Major transitions in the evolution of early land plants: a bryological perspective. Ann Bot **109**: 851–871

Lin Y, Wang K, Li X, Sun C, Yin R, Wang Y, Wang Y, Zhang M (2018) Evolution, functional differentiation, and co-expression of the RLK gene family revealed in Jilin ginseng, Panax ginseng C.A. Meyer. Mol Genet Genomics **293**: 845–859

Liu PL, Du L, Huang Y, Gao SM, Yu M (2017) Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. BMC Evol Biol **17**: 47

Liu PL, Xie LL, Li PW, Mao JF, Liu H, Gao SM, Shi PH, Gong JQ (2016) Duplication and divergence of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in basal angiosperm Amborella trichopoda. Front Plant Science **7**: 1952

Liu XS, Liang CC, Hou SG, Wang X, Chen DH, Shen JL, Zhang W, Wang M (2020) The LRR-RLK protein HSL3 regulates stomatal closure and the drought stress response by modulating hydrogen peroxide homeostasis. Front Plant Sci **11**: 548034

Lori M, van Verk MC, Hander T, Schatowitz H, Klauser D, Flury P, Gehring CA, Boller T, Bartels S (2015) Evolutionary divergence of the plant elicitor peptides (Peps) and their receptors: interfamily incompatibility of perception but compatibility of downstream signaling. J Exp Bot **66**: 5315–5325

Luo M, Dennis ES, Berger F, Peacock WJ, Chaudhury A (2005) MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. Proc Natl Acad Sci U S A **102**: 17531–17536

Ma X, Xu G, He P, Shan L (2016) SERKing coreceptors for receptors. Trends in Plant Science **21**: 1017–1033

Magalhães DM, Scholte LL, Silva NV, Oliveira GC, Zipfel C, Takita MA, De Souza AA (2016) LRR-RLK family from two Citrus species: genome-wide identification and evolutionary aspects. BMC Genomics **17**: 623

Man J, Gallagher JP, Bartlett M (2020) Structural evolution drives diversification of the large LRR-RLK gene family. New Phytol **226**: 1492–1505

Matsubayashi Y (2011) Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol **52**: 5–13

Matsubayashi Y (2014) Posttranslationally modified small-peptide signals in plants. Annu Rev Plant Biol **65**: 385–413

Matsubayashi Y, Sakagami Y (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of Asparagus officinalis L. Proc Natl Acad Sci **93**: 7623–7627

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ (2018) The timescale of early land plant evolution. Proc Natl Acad Sci U S A **115**: E2274–e2283

Muschietti JP, Wengier DL (2018) How many receptor-like kinases are required to operate a pollen tube. Curr Opin Plant Biol **41**: 73–82

Nakayama T, Shinohara H, Tanaka M, Baba K, Ogawa-Ohnishi M, Matsubayashi Y (2017) A peptide hormone required for Casparian strip diffusion barrier formation in Arabidopsis roots. Science **355**: 284–286

Nguyen L, Schmidt Heiko A, von Haeseler A, Minh B (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology And Evolution **32**: 268–274

Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, Haas FB, Vanderstraeten L, Becker D, Lang D, et al. (2018) The Chara Genome: secondary complexity and implications for plant terrestrialization. Cell **174**: 448–464.e424

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, et al. (2013) The Norway spruce genome sequence and conifer genome evolution. Nature **497**: 579–584

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y (2008) Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. Science (New York, N.Y.) **319**:294. 10.1126/science.1150083. 18202283.

Oh E, Seo PJ, Kim J (2018) Signaling peptides and receptors coordinating plant root development. Trends Plant Sci **23**: 337–351

Okuda S, Fujita S, Moretti A, Hohmann U, Doblas VG, Ma Y, Pfister A, Brandt B, Geldner N, Hothorn M (2020) Molecular mechanism for the recognition of sequence-divergent CIF peptides by the plant receptor kinases GSO1/SGN3 and GSO2. Proc Natl Acad Sci U S A **117**: 2693–2703

Olsson V, Joos L, Zhu S, Gevaert K, Butenko MA, De Smet I (2019) Look closely, the beautiful may be small: precursor-derived peptides in plants. Annu Rev Plant Biol **70**: 153–186

One Thousand Plant Transcriptomes Initiative (2019) One thousand plant transcriptomes and the phylogenomics of green plants. Nature **574**: 679–685

Pearce G, Strydom D, Johnson S, Ryan CA (1991) A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. Science **253**: 895–897

Plisson F, Ramírez-Sánchez O, Martínez-Hernández C (2020) Machine learning-guided discovery and design of non-hemolytic peptides. Sci Rep **10**: 16581

Qian P, Song W, Yokoo T, Minobe A, Wang G, Ishida T, Sawa S, Chai J, Kakimoto T (2018) The CLE9/10 secretory peptide regulates stomatal and vascular development through distinct receptors. Nat Plants **4**: 1071–1081

Racolta A, Bryan AC, Tax FE (2014) The receptor-like kinases GSO1 and GSO2 together regulate root growth in Arabidopsis through control of cell division and cell fate specification. Dev Dyn **243**: 257–278

Raven JA, Edwards D (2001) Roots: evolutionary origins and biogeochemical significance. J Exp Bot **52**: 381–401

Rensing SA (2017) Why we need more non-seed plant models. New Phytol **216**: 355–360

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, et al. (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science **319**: 64–69

Roberts I, Smith S, De Rybel B, Van Den Broeke J, Smet W, De Cokere S, Mispelaere M, De Smet I, Beeckman T (2013) The CEP family in land plants: evolutionary analyses, expression studies, and role in Arabidopsis shoot development. Journal of Experimental Botany **64**: 5371–5381

Safaeizadeh M, Boller T (2019) Differential and tissue-specific activation pattern of the AtPROPEP and AtPEPR genes in response to biotic and abiotic stress in Arabidopsis thaliana. Plant Signal Behav **14**: e1590094

Santiago J, Brandt B, Wildhagen M, Hohmann U, Hothorn LA, Butenko MA, Hothorn M (2016) Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. eLife **5**: e15075

Sauter M (2015) Phytosulfokine peptide signaling. J Exp Bot **66**: 5161–5169

Segonzac C, Monaghan J (2019) Modulation of plant innate immune signaling by small peptides. Curr Opin Plant Biol **51**: 22–28

Serrano A (2020) DeepMSPeptide: peptide detectability prediction using deep learning. Bioinformatics **36**: 1279–1280

Shi C-L, Alling RM, Hammerstad M, Aalen RB (2019) Control of organ abscission and other cell separation processes by evolutionary conserved peptide signaling. Plants (Basel) **8**: 225

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol **16**: 1114–1116

Shinohara H, Mori A, Yasue N, Sumida K, Matsubayashi Y (2016) Identification of three LRR-RKs involved in perception of root meristem growth factor in Arabidopsis. Proc Natl Acad Sci U S A **113**: 3897–3902

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

Smakowska-Luzan E, Mott GA, Parys K, Stegmann M, Howton TC, Layeghifard M, Neuhold J, Lehner A, Kong J, Grünwald K, et al. (2018) An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. Nature **553**: 342–346

Song W, Liu L, Wang J, Wu Z, Zhang H, Tang J, Lin G, Wang Y, Wen X, Li W, et al. (2016) Signature motif-guided identification of receptors for peptide hormones essential for root meristem growth. Cell Res **26**: 674–685

Steemans P, Herisse AL, Melvin J, Miller MA, Paris F, Verniers J, Wellman CH (2009) Origin and radiation of the earliest vascular land plants. Science **324**: 353

Stenvik G-E., Tandstad NM, Guo Y, Shi C-L., Kristiansen W, Holmgren A, Clark SE, Aalen RB, Butenko MA (2008a) The EPIP peptide of INFLORESCENCE DEFICIENT IN ABSCISSION Is sufficient to induce abscission in Arabidopsis through the receptor-Like kinases HAESA and HAESA-LIKE2. Plant Cell **20**: 1805–1817

Stenvik GE, Tandstad NM, Guo Y, Shi CL, Kristiansen W, Holmgren A, Clark SE, Aalen RB, Butenko MA (2008b) The EPIP peptide of INFLORESCENCE DEFICIENT IN ABSCISSION is sufficient to induce abscission in Arabidopsis through the receptor-like kinases HAESA and HAESA-LIKE2. Plant Cell **20**: 1805–1817

Stø, I.M., Orr RJ, Fooyontphanich K, Jin X, Knutsen JM, Fischer U, Tranbarger TJ, Nordal I, Aalen RB (2015) Conservation of the abscission signaling peptide IDA during Angiosperm evolution: withstanding genome duplications and gain and loss of the receptors HAE/HSL2. Front Plant Sci **6**: 931

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H, Matsubayashi Y (2014) Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. Science **346**: 343–346

Taleski M, Imin N, Djordjevic MA (2018) CEP peptide hormones: key players in orchestrating nitrogen-demand signaling, root nodulation, and lateral root development. J Exp Bot **69**: 1829–1836

Tang J, Han Z, Sun Y, Zhang H, Gong X, Chai J (2015) Structural basis for recognition of an endogenous peptide by the plant receptor kinase PEPR1. Cell Res **25**: 110–120

Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue BPA. (2015) The plant peptidome: an expanding repertoire of structural features and biological functions. Plant Cell **27**: 2095–2118

Toyokura K, Goh T, Shinohara H, Shinoda A, Kondo Y, Okamoto Y, Uehara T, Fujimoto K, Okushima Y, Ikeyama Y, et al. (2019) Lateral inhibition by a peptide hormone-receptor cascade during arabidopsis lateral root founder cell formation. Dev Cell **48**: 64–75.e65

Tsuwamoto R, Fukuoka H, Takahata Y (2008) GASSHO1 and GASSHO2 encoding a putative leucine-rich repeat transmembrane-type receptor kinase are essential for the normal development of the epidermal surface in Arabidopsis embryos. Plant J **54**: 30–42

Vanneste K, Baele G, Maere S, Van de Peer Y (2014) Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary. Genome Res **24**: 1334–1347

Vie AK, Najafi J, Liu B, Winge P, Butenko MA, Hornslien KS, Kumpf R, Aalen RB, Bones AM, Brembu T (2015) The IDA/IDA-LIKE and PIP/PIP-LIKE gene families in Arabidopsis: phylogenetic relationship, expression patterns, and transcriptional effect of the PIPL3 peptide. J Exp Bot **66**: 5351–5365

Wei Z, Wang J, Yang S, Song Y (2015) Identification and expression analysis of the LRR-RLK gene family in tomato (Solanum lycopersicum) Heinz 1706. Genome **58**: 121–134

Whitewoods CD, Cammarata J, Nemec Venza Z, Sang S, Crook AD, Aoyama T, Wang XY, Waller M, Kamisugi Y, Cuming AC, et al. (2018) CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants. Curr Biol **28**: 2365–2376.e2365

Wu H, Zheng R, Hao Z, Meng Y, Weng Y, Zhou X, Zhu L, Hu X, Wang G, Shi J, et al. (2019) Cunninghamia lanceolata PSK peptide hormone genes promote primary root growth and adventitious root formation. Plants (Basel) **8**: 520

Yang Y, Sun P, Lv L, Wang D, Ru D, Li Y, Ma T, Zhang L, Shen X, Meng F, et al. (2020) Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. Nat Plants **6**: 215–222

Zan Y, Ji Y, Zhang Y, Yang S, Song Y, Wang J (2013**)** Genome-wide identification, characterization and expression analysis of populus leucine-rich repeat receptor-like protein kinase genes. BMC Genomics **14**: 318

Zhang J, Fu XX, Li RQ, Zhao X, Liu Y, Li MH, Zwaenepoel A, Ma H, Goffinet B, Guan YL, et al. (2020a) The hornwort genome and early land plant evolution. Nat Plants **6**: 107–118

Zhang Z, Liu L, Kucukoglu M, Tian D, Larkin RM, Shi X, Zheng B (2020b) Predicting and clustering plant CLE genes with a new method developed specifically for short amino acid sequences. BMC Genomics **21**: 709

Zhou F, Guo Y, Qiu LJ (2016) Genome-wide identification and evolutionary analysis of leucine-rich repeat receptor-like protein kinase genes in soybean. BMC Plant Biol **16**: 58