Journal of Experimental Botany, Vol. 60, No. 15, pp. 4383–4396, 2009
doi:10.1093/jxb/erp277 Advance Access publication 22 September, 2009
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Arabidopsis L-type lectin receptor kinases: phylogeny, classification, and expression profiles

Klaas Bouwmeester and Francine Govers*

Laboratory of Phytopathology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen and Graduate School Experimental Plant Sciences, The Netherlands

Received 14 July 2009; Revised 8 August 2009; Accepted 17 August 2009

Abstract

In plants, lectin receptor kinases are considered to play crucial roles during development and in the adaptive response to various stimuli. *Arabidopsis* lectin receptor kinases can be divided into three type-classes based on sequence similarity of their extracellular lectin motifs. The current study focuses on the legume-like lectin receptor kinases (LecRKs), which are regarded as ideal candidates for monitoring cell wall integrity and are possibly functional in adaptive responses. An inventory of the Arabidopsis *LecRK* gene family is presented here. It consists of 45 members including three that were recently identified; two encode N-terminal truncated variants one of which has two in tandem kinase domains. Phylogenetic trees derived from full-length amino acid sequence alignments were highly concordant to phylograms that were purely based on lectin motifs or kinase domains. The phylograms allowed reclassification of the *LecRK* genes and hence a new proposal for gene nomenclature was suggested. In addition, a comprehensive expression analysis was executed by exploring public repositories. This revealed that several *LecRK* genes are differentially expressed during plant growth and development. Moreover, multiple *LecRK*s appear to be induced upon treatment with elicitors and pathogen infection. Variation in gene expression was also analysed in seedlings of diverse Arabidopsis accessions. Taken together, this study provides a genome-wide overview of the *LecRK* gene family and an up-to-date classification using a novel and systematic gene nomenclature.

Key words: *Arabidopsis*, evolutionary relationship, *LecRK*s, nomenclature, RLK.

Introduction

Plants are facing diverse developmental and environmental stimuli. Perception and transduction of these signals is largely governed by so-called receptor-like kinases (RLKs). The *Arabidopsis* genome harbours over 600 *RLK* genes, many of which can be grouped into distinct subfamilies based on their extracellular domains (Shiu and Bleecker, 2001*a*, 2003). The variability in the extracellular domain organization reflects their diverse function and mode of signal perception. For most RLKs, however, neither their ligands nor their downstream targets are known. Examples of elucidated ligand–receptor pairs include the pattern recognition receptors FLS2 and EFR that interact with the microbial associated molecular patterns (MAMPs) flg22 and EF-Tu, respectively, as well as the plant development-associated receptors CLV1, which has the polypeptide CLV3 as ligand, and BRI1 that functions in brassinosteroid perception (Gómez-Gómez and Boller, 2000; Kinoshita *et al.*, 2005; Zipfel *et al.*, 2006; Ogawa *et al.*, 2008).

The adaptive responses of plants to extracellular ligands and stimuli are governed by a functional continuum between the plant cell wall (CW) and the plasma membrane (PM) in which RLKs with CW-bound extracellular domains

* To whom correspondence should be addressed: E-mail: francine.govers@wur.nl

Abbreviations: RLK, receptor-like kinase; MAMP, microbe-associated molecular pattern; CW, cell wall; PM, plasma membrane; WAK, wall-associated kinase; PERK, proline-rich extensin-like receptor kinase; GNA, *Galanthus nivalis* agglutinin; SI, self-incompatibility; SRK, *S*-locus receptor kinase; EGF, epidermal growth factor; PAN, plasminogen-apple-nematode; LecRK, legume-like lectin receptor kinase; RGD, arginine-glycine-aspartic acid; JAK, Janus kinase; MPSS, massively parallel signature sequencing; LPS, lipo-polysaccharides; NN, neural networks; HMM, hidden Markov models.

probably function as cell wall integrity sensors (reviewed by Humphrey *et al.*, 2007). The *Arabidopsis* wall-associated kinases (WAKs) are thought to operate as physical CW-PM linkers (Anderson *et al.*, 2001; Verica and He, 2002). By means of various biochemical techniques it was shown that extracellular subdomains of WAK1 interact with a secreted glycine-rich protein (AtGRP3) and with cell wall pectins (Park *et al.*, 2001; Decreux and Messiaen, 2005). *Arabidopsis* plants with impaired *WAK4* expression were found to be affected in cell elongation and morphology (Lally *et al.*, 2001; Kohorn *et al.*, 2006*a*). Moreover, inactivation of *WAK2*, one of the *Arabidopsis WAK* genes, reduced the activity of vacuolar invertase that functions in controlling cell turgor, and hence seems to regulate cell expansion (Kohorn *et al.*, 2006*b*). Other *Arabidopsis* RLKs that may have a role in binding cell wall components are the PERKs and members of the CrRLK1L family closely resembling a RLK from *Catharanthus roseus* (Madagascar periwinkle) (Schulze-Muth *et al.*, 1996). PERKs harbour extracellular proline-rich sequences that share similarity with plant extensins, suggesting a direct interaction with cell wall components (Nakhamchik *et al.*, 2004), whereas THESEUS1 – a CrRLK1L protein from *Arabidopsis* – is activated in mutants deficient in cellulose and triggers growth inhibition and ectopic lignin accumulation (Hématy *et al.*, 2007).

Another group of RLKs that are regarded as potential CW-PM linkers are the lectin receptor kinases which contain extracellular lectin motifs that are known to bind various carbohydrates. Based on the extracellular lectin motifs three types can be distinguished; G, C, and L (Fig. 1). *Arabidopsis* has around 40 *RLK* genes encoding bulb-associated lectin receptor kinases. Previously these were categorized as B-type lectin receptor kinases because their lectin motif resembles mannose-specific lectins of bulb species. Because these lectins are no longer confined to bulbs Van Damme *et al.* (2007) proposed to rename them as GNA-related lectins, here abbreviated as G-type lectins. Well-studied G-type lectin receptor kinases are the *S*-locus receptor kinases, historically known as *S*-domain RLKs after the functionality of the ectodomain in self-incompatibility (SI) in flowering plants. In Brassicaceae, the *S*-locus receptor kinase SRK acts as the stigmatic determinant of the SI response and possibly functions as a receptor for the pollen ligand SCR/SP11. The shift to self-fertility in *Arabidopsis thaliana* is partly due to non-functional SRK sequences (Kusaba *et al.*, 2001; Sherman-Broyles *et al.*, 2007). G-type lectin receptor kinases were also shown to play roles in plant defence. The rice gene *Pi-d2*, for example, confers resistance to the fungal pathogen *Magnaporthe grisea* whereas NgRLK1 from *Nicotiana glutinosa* was selected in a yeast two-hybrid screen as a putative interactor with capsicein, an elicitin from *Phytophthora capsici* (Chen *et al.*, 2006; Kim *et al.*, 2009). Besides the G-type lectin motifs the extracellular domains of these proteins contain cysteine-rich EGF-like (epidermal growth factor) and PAN (plasminogen-apple-nematode) motifs that both function in protein homodimerization (Naithani *et al.*, 2007). As yet, the role of the G-lectin motif is unknown and there is also no evidence for a function in ligand binding.

C-type (calcium-dependent) lectin motifs can be found in a large number of mammalian proteins that mediate innate immune responses such as antigen uptake and T-cell interaction and play a major role in pathogen recognition (reviewed by Cambi *et al.*, 2005). Unlike the omnipresent nature of C-type lectins in mammals, they seem to be scarce in plants. *Arabidopsis* has only a single gene encoding a protein with a C-type lectin motif but so far its function has not been elucidated.

The third type comprises the legume-like or L-type lectin receptor kinases (LecRKs). Their extracellular domains resemble soluble legume lectins which are ubiquitous in leguminous seeds. For several legume lectins the 3D-structure and their carbohydrate-protein binding specificity has been elucidated. The sequence similarity to legume lectins prompted the hypothesis that the potential ligands of LecRKs are oligosaccharides (André *et al.*, 2005). Molecular modelling showed that the sugar-binding residues in LecRKs are poorly conserved and it is therefore unlikely that these receptors bind monosaccharide molecules. There

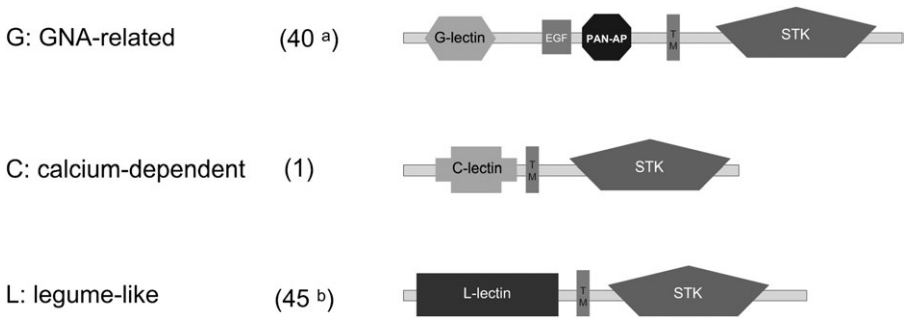

Fig. 1. Domain composition and organization of the lectin receptor kinases. Based on the lectin motifs in the extracellular domain, lectin receptor kinases are divided in three types, G, C, and L. The numbers in brackets refer to the number of genes in *Arabidopsis* encoding G, C, and L-type lectin receptor kinases. G-type lectin receptor kinases, which contain GNA-related lectin motifs, were previously categorized as B-type. [a] Number as listed by Shiu and Bleecker (2001*b*); [b] Includes the two lectin receptor kinases with a N-truncated terminus. TM, transmembrane; PAN, plasminogen-apple-nematode motif; EGF, epidermal growth factor-like motif; STK, serine/threonine protein kinase domain.

is, however, a rather conserved hydrophobic-binding site, and hence LecRKs may serve in the recognition of small hydrophobic ligands, such as plant hormones or MAMPs (Barre *et al.*, 2002; André *et al.*, 2005).

Although LecRKs are implied to function in diverse biological processes, their exact biological role has not yet been clarified. LecRKs that have been described to function during plant development include the SGC lectin RLK of *Arabidopsis*, which is required for proper pollen development (Wan *et al.*, 2008) and MtLecRK1;1 from the nitrogen-fixing legume *Medicago truncatula* that influences nodulation (Navarro-Gochicoa *et al.*, 2003). Expression of the *Arabidopsis* lectin receptor kinase *lecRK-a1* was shown to be induced during wounding, senescence of leaves, and in response to oligogalacturonides, which might be released upon disruption of the plant cell wall (Riou *et al.*, 2002). More recently, Xin *et al.* (2009) showed that a specific subfamily of LecRKs is responsible for negatively regulating the abscisic acid (ABA) response during seed germination and hypothesized that these genes directly or indirectly affect defence. Lately, several reports have presented data linking ABA-signalling with defence responses (Asselbergh *et al.*, 2008). NbLRK1, a legume-like lectin receptor kinase from *Nicotiana benthamiana* was reported to interact intracellularly with the *P. infestans* elicitin INF1 and seems to be involved in the subsequent INF1-induced cell death (Kanzaki *et al.*, 2008). Another LecRK with a potential link to plant defence and pathogen response is LecRK79 in *Arabidopsis*. It interacts via its lectin motif with the tripeptide motif RGD (arginine-glycine-aspartic acid) in IPI-O, an RXLR effector from the late blight pathogen *Phytophthora infestans* that can disrupt CW-PM adhesions through the RGD motif (Senchou *et al.*, 2004; Gouget *et al.*, 2006). LecRK79 mediates CW-PM adhesions and hence the continuum between the cell wall and the plasma membrane (Gouget *et al.*, 2006). Its role in plant defence is, furthermore, supported by the observation that *LecRK79* expression is induced upon inoculation with several non-host and avirulent pathogens (Bouwmeester *et al.*, 2008). Taken together, these data suggests that LecRKs play crucial roles in both developmental and adaptive processes.

Over the past years several *Arabidopsis LecRK*s have been investigated. Gene naming, however, has not been uniform and this has sometimes resulted in confusion and miscommunication. Some *LecRK* genes have multiple names, whereas the majority has only an AGI gene code. Gene names are often based on protein functionality or molecular mass, but this does not reflect their phylogenetic relationship. Unfortunately, several gene names have been chosen based on incomplete and not always informative gene classification. For example, *LecRK-a4* – named by Hervé *et al.* (1996) – strongly resembles *LecRKA4.1*, the gene name recently assigned by Xin *et al.* (2009) to a LecRK that regulates the ABA response.

In this study, the information available on all members of the *Arabidopsis* L-type lectin receptor kinase gene family has been collated and summarized and a bioinformatic analysis has been performed to clarify their relationship. Based on the revised phylogenetic classification, a simplified nomenclature for the *Arabidopsis LecRK*s is proposed, which could serve as a basis for gene naming in other plant species and will hopefully improve communication among scientists working on this group of plant receptor kinases. A comprehensive survey of publicly available expression data has also been undertaken. The expression profiles show that *LecRK*s are activated by various biotic and abiotic stimuli and are differentially expressed in various *Arabidopsis* accessions.

Materials and methods

Sequence database search and gene analysis

Gene sequences were retrieved using the available servers at the TAIR website (http://www.arabidopsis.org). Additional BLAST searches were conducted on the *Arabidopsis* genome sequence and proteome at TAIR and MIPS (http://mips.gsf.de). All genes were checked for annotation mistakes. Structural domain prediction was performed using the SMART (http://smart.embl-heidelberg.de) and Pfam (http://pfam.sanger.ac.uk) databases.

SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP) was used for the prediction of signal peptides. The web servers TMHMM (http://www.cbs.dtu.dk/services/TMHMM), SOSUI (http://bp.nuap.nagoya-u.ac.jp/~sosui/), and PRED-TMR2 (http://athina.biol.uoa.gr/PRED-TMR2/) were used to predict transmembrane regions. Additional domain prediction and verification was performed using the programs InterProScan (http://www.ebi.ac.uk/Tools/InterProScan/) and ProDom (http://prodom.prabi.fr). Homology at protein and nucleotide levels was determined by pairwise sequence comparison with the AlignX module in Vector NTI® (Invitrogen).

Alignment and phylogenetic analysis of LecRK *genes*

Full-length protein and domain sequences, as defined in the previous section, were compiled and aligned using ClustalW (Larkin *et al.*, 2007). GONNET with a penalty gap opening of 10 and a gap extension of 0.1 was used as the protein weight matrix. The obtained sequence alignments were used as input to construct phylogenetic trees with the minimal evolution algorithm within MEGA 4 (Tamura *et al.*, 2007). Poisson correction was used to account for multiple substitutions and alignment gaps were removed in a pairwise fashion. The statistical strength was assessed by bootstrap resampling using 10 000 replicates.

Analyses conducted with different algorithms within MEGA 4 generated comparable phylograms. Protein phylogenetic trees were rooted with amino acid sequences of the *Arabidopsis* RLKs WAK1 (At1g21250), PERK1 (At3g24550), the G-type lectin receptor kinases (SRKs) ARK1 (At1g65790) and CES101 (At3g16030), and the C-type lectin receptor kinase encoded by the single-copy gene At1g52310. Domain-based phylogenetic trees were

rooted with the predicted extra- and intracellular domains of these outgroup genes. Branches corresponding to partitions reproduced in <50% bootstrap replicates were collapsed.

Analysis of expression data

Arabidopsis EST and cDNA data was retrieved from the TAIR website. MPSS signature data were retrieved from the MPSS database (http://mpss.udel.edu/at/) (Lu *et al.*, 2005). Sums of abundance of sense strand signatures (classes 1, 2, 5, and 7) were regarded as significant and reliable MPSS expression data. Sequence queries in the 17 bp and 20 bp signature databases showed comparable expression profiles. Expression data was further analysed by using the Genevestigator V3 web tool (http://www.genevestigator.com) (Hruz *et al.*, 2008). Expression analysis was restricted to gene chip experiments based on the sequenced genome of the *Arabidopsis* accession Col-0. *LecRK-V.7* (At3g58740) and *LecRK-V.8* (At3g59750) share the same probe-set (25148_s_at) and represent equal values. The blue colour codes displayed in Fig. 3 are based on the concept of expression potential, in which the darkest blue colour is representing the maximum expression level that a given gene can reach throughout all arrays available in the database. Because equal blue colours can depict different expression values, comparison of expression levels is only justified between one gene in different tissues and not between individual genes.

Natural variation in *LecRK* expression between *Arabidopsis* accessions was determined by using the *Arabidopsis* eFP browser of the Bio-Array Resource (BAR) website (http://bar.utoronto.ca/) (Winter *et al.*, 2007). Log_2 expression values relative to the control value (Col-0) were tabulated and reformatted to a colour-scale with the BAR HeatMapper Plus Tool to visualize expression differences easily. Expression was measured in triplicate samples of aerial parts of 4-d-old seedlings of nine geographically separated *Arabidopsis* accessions; including the reference accession Col-0 (Columbia, USA), Van-0 (Vancouver, USA), Kin-0 (Kindalville, USA), Est-0 (Estonia), Bay-0, (Bayreuth, Germany), Nd-0 (Niederzenz, Germany), L*er*-2 (Landsberg, Germany), Sha (Shahdara, Tajikistan), and Cvi (Cape Verde Islands).

Results and discussion

Gene identification and protein structure prediction

L-type lectin receptor kinases (LecRKs) were identified by BLAST searches and by exploring previously constructed gene lists (Shiu and Bleecker, 2001*b*, 2003; Barre *et al.*, 2002). In total, 45 *LecRK*s were identified, 42 of which were previously described by Barre *et al.* (2002) (Table 1). The three added in this survey are At5g60310, At3g46760, and At2g32800, which were previously listed as putative lectin receptor kinases in the *Arabidopsis* RLK inventory by Shiu and Bleecker (2001*b*, 2003). At5g60310 belongs to a cluster of three closely linked *LecRK* genes, and seems to encode a typical L-type lectin receptor kinase (Fig. 1). Protein sequence comparison revealed, however, that At5g60310 encodes an abnormal protein kinase lacking sub-domains X and XI, thus suggesting impaired kinase activity. The other two, At3g46760 and At2g32800, encode proteins with kinase domains but lacking transmembrane motifs and extracellular domains. Moreover, the kinase encoded by At3g46760 contains a shortened activation loop, which could affect proper kinase functioning. Structure prediction of At2g32800 revealed a peculiar arrangement of two repeated kinase domains (here named AT3G32800A and AT3G32800B) that contain all structural sub-domains. Proteins with two in-tandem kinase domains do exist in vertebrates and are known as Janus kinases (JAKs) based on their two-faced function: the first domain exhibits kinase activity while the other down-regulates this activity. JAKs appear to play a critical role in diverse processes such as cytokinin signalling in mammalian cells (reviewed by Yamaoka *et al.*, 2004).

Prediction analyses showed that most LecRKs have all the features of typical L-type lectin receptors, i.e. a putative signal sequence, an extracellular lectin motif, a transmembrane region, and a protein kinase domain (Fig. 1). In addition to the three exceptions described above, four other *LecRK*s encode structurally different proteins. The one encoded by At4g28350 was predicted to lack a transmembrane region, and those encoded by At2g29220, At2g29250, and At3g45390 seem to be devoid of a typical kinase structure (Table 1).

LecRK *clusters and clades*

To categorize the *Arabidopsis LecRK* genes, the *LecRK* gene family was divided into clusters and clades. A gene cluster represents physical proximity on the chromosome and was defined as two or more homologous genes within a maximum of eight adjacent gene models between individual members. Nine distinct *LecRK* gene clusters were found with a maximum of six genes per cluster (i.e. cluster 3A) (Table 1). The gene clusters are dispersed across the five *Arabidopsis* chromosomes, but the larger gene clusters are predominantly located on chromosome 3 and 5.

A gene clade was defined as a group of at least two homologous genes with a minimum of 50% identity at both the nucleotide and amino acid level (Mondragon-Palomino and Gaut, 2005). By means of pairwise comparison nine major *LecRK* clades were determined. The two largest *LecRK* clades comprise 11 and nine genes, respectively. Two clades contain four genes, and the five smallest ones consist of two *LecRK* genes. Seven *LecRK* genes do not belong to any clade and were assigned as singletons. The 28 *LecRK* genes distributed over clusters all belong to a clade. Hence, physically linked *LecRK* genes also contain a high pairwise identity (>50%). This is probably due to ancient duplication events and it is therefore not surprising that genes in one cluster belong to a single clade (Table 1).

Table 1. *LecRK* classification, nomenclature, and coding sequence information

Proposed clade name	Proposed gene name	Cluster[a]	Locus[b]	Alternative gene name	Former class[c]	Gene alias[b]	Gene models	Protein accession[d]	SP NN/HMM[e]	No. of TM motifs[f]	Protein length (AA)	Remarks
LecRK-I	*LecRK-I.1*	3A	At3g45330		C	F18N11.90	1	Q9M3E5	y/y	2	682	
	LecRK-I.2	3A	At3g45390		C	F18N11.150	1	Q7FK82	y/y	1	604	lacks kinase sub-domain V, VIa, VIb
	LecRK-I.3	3A	At3g45410	*LecRK2*	C	F18N11.170	1	Q9M3D8	y/y	1	664	[g]
	LecRK-I.4	3A	At3g45420		C	F18N11.180	1	Q9M3D7	y/y	1	667	
	LecRK-I.5	3A	At3g45430		C	F9K21.10	1	Q9M1G4	y/y	2	613	annotated as lacking SP
	LecRK-I.6	3A	At3g45440		C	F9K21.20	1	Q9M1G3	y/y	2	669	
	LecRK-I.7	5A	At5g60270		C	F15L12.9	1	Q9LSS0	y/y	1	668	
	LecRK-I.8	5A	At5g60280		C	F15L12.12	1	Q9LSR9	y/y	1	657	
	LecRK-I.9	5A	At5g60300	*LecRK79*	C	F15L12.17	3	Q56XH0	y/y	1	718/766	[h]
	LecRK-I.10	5A	At5g60310			F15L12.19	1	Q3E884	y/n	1	616	lacks kinase sub-domain X and XI
	LecRK-I.11	5A	At5g60320		C	K9B18.1	1	Q9FJI4	y/n	2	675	
LecRK-II	*LecRK-II.1*	5B	At5g59260		C	MNC17.17	1	Q9FIF1	y/n	1	674	
	LecRK-II.2	5B	At5g59270		C	MNC17.20	1	absent	y/y	1	668	
LecRK-III	*LecRK-III.1*	2A	At2g29220			F16P2.40	1	Q9ZW09	y/y	1	627	incomplete activation loop
	LecRK-III.2	2A	At2g29250			F16P2.37	1	Q9ZW11	y/y	1	623	incomplete activation loop
LecRK-IV	*LecRK-IV.1*	-	At2g37710	*LRK1*, *LecRK-d*	A2	F13M22.21	1	O80939	y/y	1	675	[i, j]
	LecRK-IV.2	-	At3g53810	*SGC lectin RLK*	A2	F5K20.110	1	Q9M345	y/y	2	677	[k]
	LecRK-IV.3	4	At4g02410		A2	T14P8.3	1	O81292	y/y	2	674	
	LecRK-IV.4	4	At4g02420		A2	T14P8.4	1	O81291	y/n	1	669	
LecRK-V	*LecRK-V.1*	1	At1g70110	*LecRK-b2*	A3	F20P5.16	1	O04534	y/y	1	666	[j]
	LecRK-V.2	1	At1g70130	*LecRK-b1*	A3	F20P5.15	1	O04533	y/y	1	656	[j]
	LecRK-V.3	2B	At2g43690	*LecRK-c1*	A1	F18O19.20	1	O22834	y/n	2	664	[j]
	LecRK-V.4	2B	At2g43700	*LecRK-c2*	A1	F18O19.19	1	O22833	y/y	1	658	[j]
	LecRK-V.5	3B	At3g59700	*Ath-lecRK1*, *LecRK1*, *lecRK-a1*	A1	T16L24.250	1	Q96285	y/n	2	661	[j, l]
	LecRK-V.6	3B	At3g59730	*LecRK-a2*	A1	T16L24.280	1	Q9LEA3	y/n	1	523	[l]
	LecRK-V.7	3B	At3g59740	*LecRK3*, *LecRK-a3*	A1	T16L24.290	1	Q9ZR79	y/y	1	659	[j, l]
	LecRK-V.8	3B	At3g59750	*LecRK-a4*	A1	F24G16.20	1	Q9M1Z9	y/n	1	626	[l]; has a frameshift mutation
	LecRK-V.9	-	At4g29050		A3	F19B15.80	1	Q9SZD5	y/y	1	669	
LecRK-VI	*LecRK-VI.1*	-	At3g08870		A4	T16O11.20	1	Q9SR87	y/n	1	693	
	LecRK-VI.2	5C	At5g01540	*LecRKA4.1*	A4	F7A7.60	1	Q9M021	y/y	1	682	[m]
	LecRK-VI.3	5C	At5g01550	*LecRKA4.2*	A4	F7A7.70	1	Q9M020	y/y	1	688	[m]
	LecRK-VI.4	5C	At5g01560	*LecRKA4.3*	A4	F7A7.80	1	Q66GN2	y/y	1	691	[m]
LecRK-VII	*LecRK-VII.1*	-	At4g04960		B3	T32N4.9	1	Q9S9U1	y/y	1	686	
	LecRK-VII.2	-	At4g28350	*LecRK-e*	B3	F20O9.40	1	O49445	y/y	0	649	[j]; no TM predicted
LecRK-VIII	*LecRK-VIII.1*	-	At3g53380		B1	F4P12.80	1	Q9LFH9	y/y	1	715	
	LecRK-VIII.2	-	At5g03140		B1	F15A17.170	1	Q9LYX1	y/n	1	711	
LecRK-IX	*LecRK-IX.1*	-	At5g10530		B2	F12B17.120	1	Q9LXA5	y/y	1	651	
	LecRK-IX.2	-	At5g65600		B2	K21L13.11	1	Q9LSL5	y/n	2	675	

Table 1. *Continued*

Proposed clade name	Proposed gene name	Cluster[a]	Locus[b]	Alternative gene name	Former class[c]	Gene alias[b]	Gene models	Protein accession[d]	SP NN/HMM[e]	No. of TM motifs[f]	Protein length (AA)	Remarks
Singletons	*LecRK-S.1*	-	At1g15530		B3	T16N11.4	1	Q9M9E0	y/y	1	656	
	LecRK-S.2	-	At2g32800	*AP4.3A*		F24L7.6	1	O48837	n/n	0	851	[n]; lacks lectin domain and TM; has two repeated kinase domains
	LecRK-S.3	-	At3g46760			T6H20.210	1	Q9STF0	n/n	0	337	lacks lectin domain and TM
	LecRK-S.4	-	At3g55550			T22E16.210	1	Q9M2S4	y/y	2	684	
	LecRK-S.4	-	At5g06740		B2	MPH15.10	1	Q9FG33	y/n	1	652	
	LecRK-S.6	-	At5g42120		B1	MJC20.23	1	Q9FHX3	y/n	1	691	
	LecRK-S.7	-	At5g55830		B1	MDF20.27	1	Q9FHG4	y/y	1	681	

[a] A dash (-) indicates that this gene does not belong to a cluster; [b] Locus and gene alias assigned on the BAC location according to TAIR; [c] gene classification as described by Barre *et al.* (2002); [d] protein designation in UniProt KB; [e] signal peptide prediction based on neural networks (NN) and hidden Markov models (HMM), y = predicted, n = not predicted; [f] prediction of transmembrane (TM) motifs based on HMMs; [g] He *et al.* (2004); [h] Gouget *et al.* (2006); [i] Swarup *et al.* (1996); [j] Hervé *et al.* (1996); [k] Wan *et al.* (2008); [l] Riou *et al.* (2002); [m] Xin *et al.* (2009); [n] Cordeiro *et al.* (1998).

LecRK *clade phylogeny*

To determine the evolutionary relationship between members of the LecRK protein family the protein sequences were aligned with ClustalW using GONNET as the protein weight matrix. Subsequently, phylograms were constructed with MEGA 4 using several algorithms, i.e. Neighbor–Joining, maximum parsimony, and minimal evolution. The robustness of the phylograms was determined by bootstrap analysis with 10 000 replications. The consensus topologies generated by the different algorithms were comparable (data not shown). All phylograms were rooted using the *Arabidopsis* RLKs WAK1, PERK1, the *Arabidopsis* C-type lectin receptor kinase (At1g52310), and the G-type lectin receptor kinases ARK1 and CES101 as outgroup representatives and were largely supported by high bootstrap values (Fig. 2).

Phylogenetic analysis conducted with 43 out of 45 full-length LecRK proteins resulted in 38 LecRKs in nine distinct clades and five singletons, which is in line with the pairwise comparisons (Fig. 2A). In this analysis At2g32800 and At3g46760 were excluded because they lack extracellular domains. The phylogram based on the lectin domain shows comparable gene relationships and also similar grouping as depicted in the phylogenetic tree generated with the full-length protein sequences (Fig. 2B). Similarly, the phylogram based on the kinase domain is comparable. As expected the three kinase domains added in the phylogram At2g32800A, At2g32800B, and At3g46760, group with the LecRKs and not with the outgroups. Moreover, they do not group with any clade and even the two kinase domains within At2g32800 behave as singletons (Fig. 2C). Overall these phylograms suggest that the LecRK members of the same clade are highly identical to each other in both extracellular and intracellular domains. This feature can be used for the phylogenetic characterization of full-length and truncated LecRK protein sequences of other plant species.

Uniform nomenclature for LecRK *gene members*

The proposed gene nomenclature is based on the grouping of closely related *LecRK* genes within the evolutionary divergent clades. Clades were designated by the Roman numerals I to IX starting clockwise within the phylogram shown in Fig. 2A. The individual gene members within the clades were given Arabic numerals based on their physical map position. Singletons were also ordered based on their physical map position and numbered S.1 to S.7. Table 1 lists the revised and more consistent gene nomenclature. For comparison, the previous gene classification (Barre *et al.*, 2002) and the alternative *LecRK* gene names are included.

LecRK *genes are differentially expressed in various tissues and developmental stages*

To obtain more insight into the expression patterns of the individual *LecRK*s during plant development, a comprehensive expression analysis was performed by making use of various on-line repositories. Based on the number of

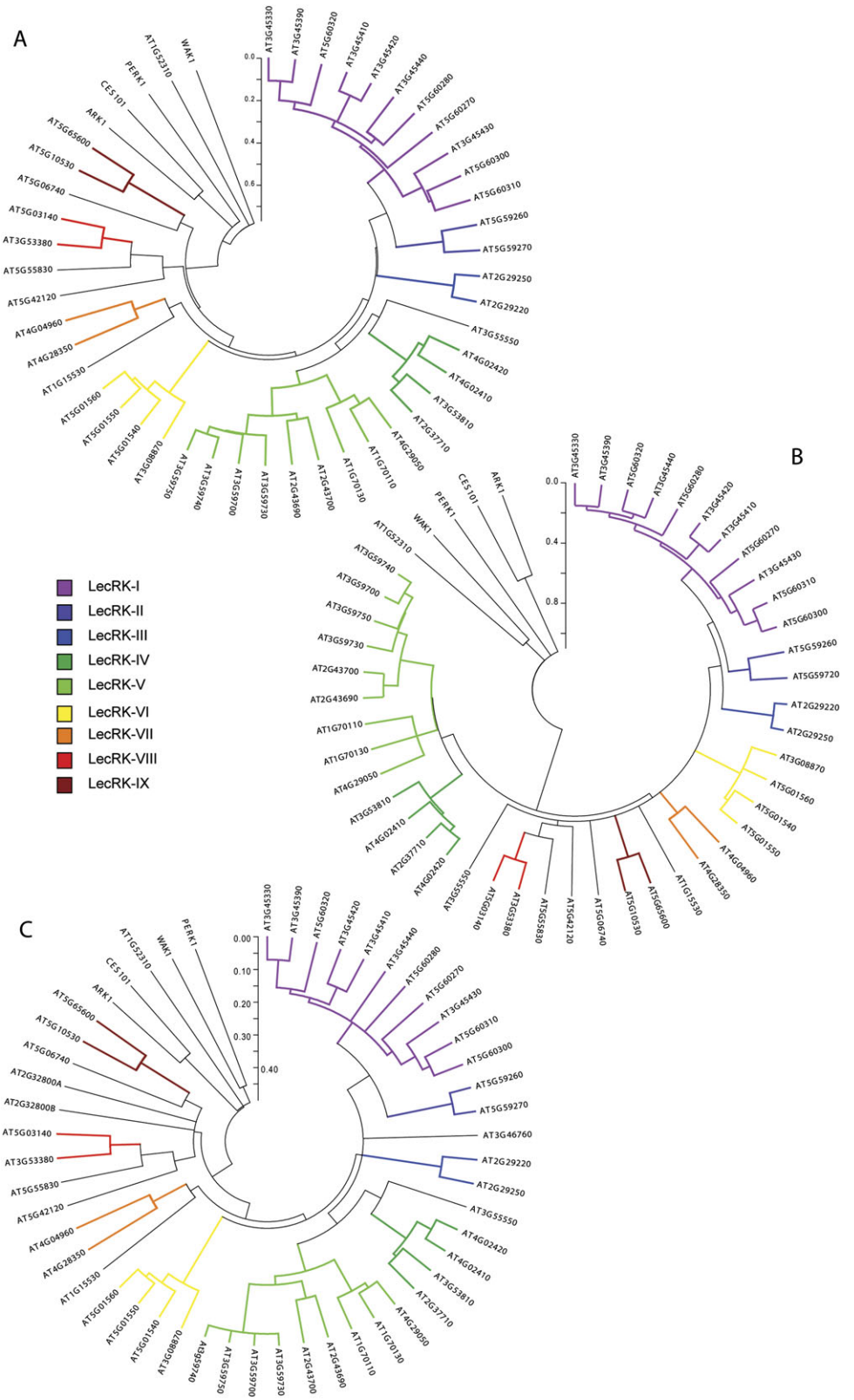

Fig. 2. Phylogenetic analysis and classification of *Arabidopsis* LecRK proteins. Phylograms of (A) 43 full-length LecRK amino acid sequences, (B) 43 lectin domains and, (C) 46 kinase domains, including the kinase domain of LecRK-S.2 and the two kinase domains of LecRK-S.3. Each LecRK clade is a depicted by a different colour.

full-length cDNAs and ESTs retrieved from the TAIR website, it is apparent that the *LecRK* genes have variable expression profiles. For about one-third of the *LecRK* genes there are no ESTs, for another one-third there are fewer than 10 ESTs, whereas five of them have over 20 ESTs and thus seem to be highly expressed (Table 2). Transcript

Table 2. cDNA, EST and MPSS 20 bp tag expression data of the *Arabidopsis LecRK* genes

Clade[a]	Gene name[a]	Locus[a]	cDNA[b]		No. ESTs[b]	No. MPSS20 bp tags[c,d,e]															
			+/-	n.		CAF	CAS	GSE	LEF	LES	ROF	ROS	AP1	AP3	AGM	SAP	INF	INS	SIF	SIS	S04	S52
LecRK-I	*LecRK-I.1*	At3g45330	-		0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0
	LecRK-I.2	At3g45390	-		0	0	0	0	0	0	5	0	0	3	0	0	0	0	0	0	0	0
	LecRK-I.3	At3g45410	+	1	7	6	0	0	8	0	54	0	0	0	0	0	0	0	0	0	0	0
	LecRK-I.4	At3g45420	-		0	0	0	0	0	0	0	0	0	9	0	0	0	2	3	0	0	0
	LecRK-I.5	At3g45430	-		3	0	0	0	2	0	9	0	0	0	0	0	45	0	0	0	0	0
	LecRK-I.6	At3g45440	-		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	LecRK-I.7	At5g60270	+	1	14	15	46	16	24	28	3	3	0	0	0	0	0	0	0	0	0	0
	LecRK-I.8	At5g60280	+	2	11	19	2	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	LecRK-I.9	At5g60300	+	5	23	64	84	58	31	37	209	82	15	8	11	38	47	15	8	12	7	0
	LecRK-I.10	At5g60310	+		1	0	9	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0
	LecRK-I.11	At5g60320	-		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LecRK-II	*LecRK-II.1*	At5g59260	-		1	0	0	0	1	0	11	0	0	0	0	0	0	0	0	0	0	0
	LecRK-II.2	At5g59270	-		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LecRK-III	*LecRK-III.1*	At2g29220	-		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	LecRK-III.2	At2g29250	-		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LecRK-IV	*LecRK-IV.1*	At2g37710	+	2	54	92	92	11	147	49	72	35	72	33	42	145	82	79	125	66	91	1690
	LecRK-IV.2	At3g53810	+	1	12	38	0	2	8	0	28	0	0	5	10	0	16	0	38	0	0	0
	LecRK-IV.3	At4g02410	+	2	13	8	0	0	49	2	12	24	15	3	3	0	11	0	0	9	0	38
	LecRK-IV.4	At4g02420	-		5	0	0	0	29	7	0	0	11	40	5	32	23	20	192	22	0	7
LecRK-V	*LecRK-V.1*	At1g70110	-		0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
	LecRK-V.2	At1g70130	-		0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0
	LecRK-V.3	At2g43690	-		0	4	6	0	0	2	7	9	2	3	0	91	0	0	0	0	0	0
	LecRK-V.4	At2g43700	+	1	3	0	0	0	1	5	0	0	0	0	0	0	4	0	0	0	0	0
	LecRK-V.5	At3g59700	+	1	6	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
	LecRK-V.6	At3g59730	-		0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
	LecRK-V.7	At3g59740	-		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	LecRK-V.8	At3g59750	-		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	LecRK-V.9	At4g29050	-		5	0	0	19	4	0	0	1	11	14	6	0	0	5	0	0	0	0
LecRK-VI	*LecRK-VI.1*	At3g08870	-		4	0	0	0	30	0	0	0	5	0	0	0	0	0	0	0	0	9
	LecRK-VI.2	At5g01540	+	2	31	0	0	4	5	7	6	18	4	13	13	9	29	5	0	0	0	91
	LecRK-VI.3	At5g01550	-		6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	LecRK-VI.4	At5g01560	+	2	36	0	0	13	0	16	0	4	0	7	2	12	12	12	0	8	0	7
LecRK-VII	*LecRK-VII.1*	At4g04960	+	2	16	29	104	3	10	6	12	15	11	13	16	14	7	0	6	0	43	0
	LecRK-VII.2	At4g28350	-		3	7	11	0	0	3	0	17	0	0	0	0	0	0	0	0	0	34
LecRK-VIII	*LecRK-VIII.1*	At3g53380	+	2	12	20	10	231	18	17	11	79	51	82	24	69	130	71	25	56	0	3
	LecRK-VIII.2	At5g03140	+	3	15	28	0	10	60	19	44	0	14	18	5	35	22	11	73	0	0	1
LecRK-IX	*LecRK-IX.1*	At5g10530	+	1	1	0	4	0	0	0	0	41	0	13	2	0	0	4	18	0	0	0
	LecRK-IX.2	At5g65600	+	1	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Singletons	*LecRK-S.1*	At1g15530	+	1	1	0	8	0	0	3	0	3	0	4	0	0	0	0	0	0	2	0
	LecRK-S.2	At2g32800	+	1		5	0	24	30	20	8	15	0	5	0	0	8	2	5	0	0	54
	LecRK-S.3	At3g46760	-		0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0
	LecRK-S.4	At3g55550	+	1	0	7	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0

Table 2. *Continued*

Clade[a]	Gene name[a]	Locus[a]	cDNA[b]		No. ESTs[b]	No. MPSS20 bp tags[c,d,e]													
			+/-	n.		CAF	CAS	GSE	LEF	LES	ROF	ROS	AP1	AP3	AGM	SAP	INF	INS	SIF	SIS	S04	S52
	LecRK-S.5	At5g06740	-		2	8	32	23	0	1	18	40	0	0	0	0	0	0	0	0	0	0
	LecRK-S.6	At5g42120	+	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	LecRK-S.7	At5g55830	-		1	0	18	0	0	0	2	0	7	9	0	29	4	4	0	0	0	0

[a] see **Table 1**; [b] cDNA and EST expression data as retrieved from TAIR; [c] sum of abundance for sense orientated classes; [d] The MPSS tag abbreviations are: CAF, callus set 1; CAS, callus set 2; GSE, germinating seedlings; LEF, leaves set 1; LES, leaves set 2; ROF, root set 1; ROS, root set 2; AP1, ap1-10 inflorescence mixed stage, immature buds; AP3, ap3-6 inflorescence mixed stage, immature buds; AGM, agamous inflorescence mixed stage, immature buds; SAP, sup/ap1 inflorescence mixed stage, immature buds; INF, inflorescence set 1; INS, inflorescence set 2; SIF, silique set 1; SIS, silique set 2; S04, salicylic acid 4 h after application; S52, salicylic acid 52 h after application; [e] MPSS 17 bp tag data are added as **Table S1**.

abundance was verified by analysing expression data generated by massively parallel signature sequencing (MPSS) (Lu *et al.*, 2005). This showed that several gene members that lack or have few ESTs do have matching expression signatures (Table 2; see Supplementary Table S1 at *JXB* online). *LecRK-V.3*, for example, has no ESTs and low MPSS values in several *Arabidopsis* tissues, but displays a specific high MPSS value in the sup/ap1 immature inflorescence stage. Similarly, *LecRK-V.6* lacks ESTs but MPSS signatures revealed a subtle expression in root tissue. Gene expression profiles based on microarray data (Fig. 3) also showed some inconsistencies with EST counts or MPSS signatures. Taken together, however, it can be concluded that several *LecRK* genes are not or hardly expressed in any of the *Arabidopsis* tissues or developmental stages, for example, several members of the LecRK-V clade and both *LecRK-III* genes (Fig. 3), whereas a few others are broadly expressed in all tissues, in particular *LecRK-I.9*, *LecRK-IV.1*, and *LecRK-VIII.1*.

LecRK *gene expression in response to various stimuli*

Microarray data were also queried to reveal changes in *LecRK* expression upon treatment with various hormones, environmental stresses, elicitors and pathogens. As shown in Fig. 4 several *LecRK* genes, including some that are not or hardly expressed during plant development, were activated or repressed in response to diverse stimuli.

Treatment of *Arabidopsis* with the hormones auxin and ethylene did not result in significant changes in expression (Fig. 4). By contrast, clear changes in transcript levels of some *LecRK* genes were detected in plants treated with ABA, brassinolide, cytokinin, methyl jasmonate, and salicylic acid (SA). MPSS signature data suggest an extremely high activity of *LecRK-IV.1* at 52 h after SA application (Table 2; see Supplementary Table S1 at *JXB* online). The same, but to a lesser extent, is true for *LecRK-VI.2*. Both *LecRK-IV.1* and *LecRK-VI.2* showed differential expression in different (untreated) tissues (Fig. 3) but it seems that this is not correlated to SA sensitivity; other *LecRK* genes with differential expression during development were not activated by SA.

In addition, several abiotic stresses influenced *LecRK* expression (Fig. 4). In response to wounding, potassium deprivation, and salt stress there were hardly any changes. Osmotic stress, anoxia, and cold, resulted in minor overall changes with the exception of one or two genes that strongly responded to one particular treatment. An example is *LecRK-V.6* which was strongly repressed by cold treatment. In the case of drought many of the *LecRK* genes responded, two of which were highly stimulated and two others strongly repressed. Response to treatment with hydrogen peroxide was particularly associated with changes in expression levels of genes in clade VII, VIII, and IX and the same genes, as well as some additional ones, were induced by ozone treatment. The most severe changes were observed in the dark when

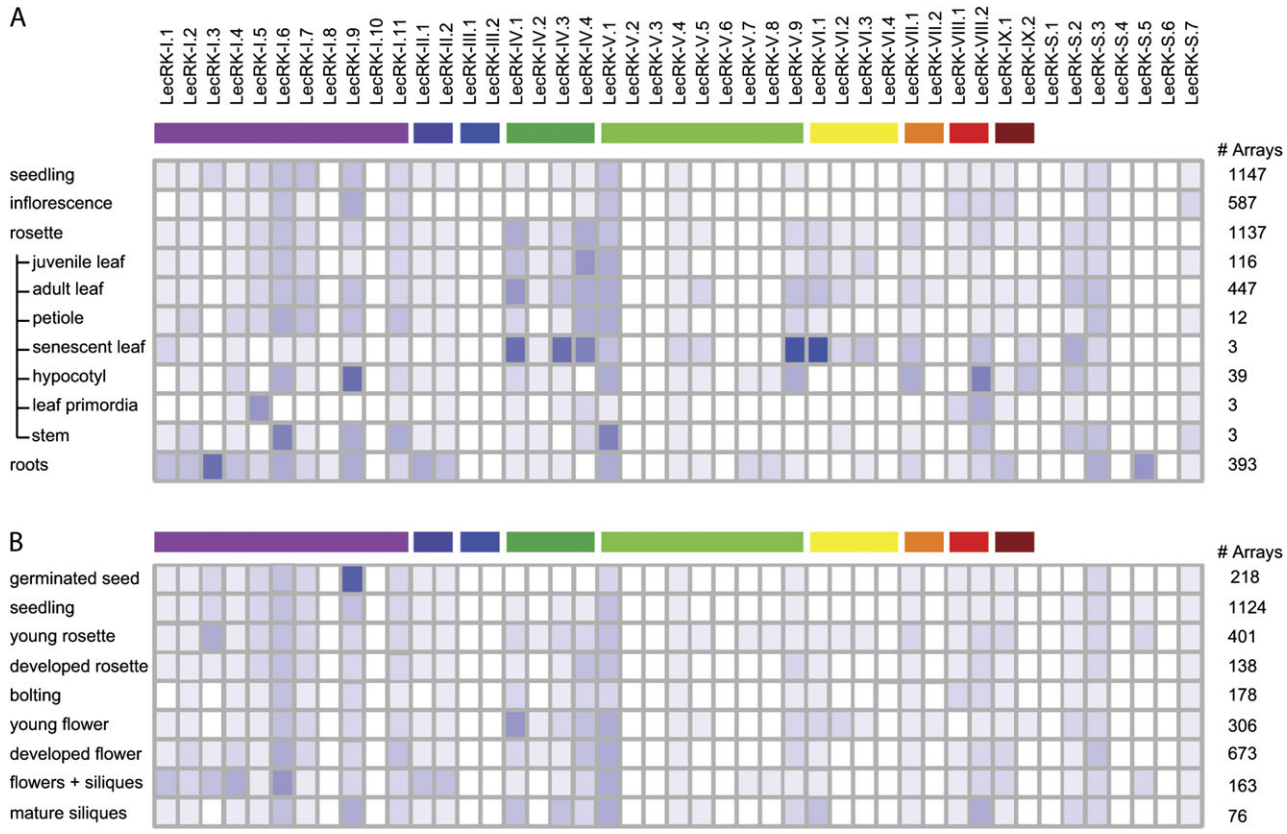

Fig. 3. Expression profiles of the *LecRK* gene family in different tissues and developmental stages of *Arabidopsis*. (A) Heat map showing levels of gene expression across various tissues. (B) Heat map showing levels of gene expression at different stages of development. The expression profile of each individual gene was normalized in such a way that the highest signal intensity is defined as 100% (dark blue) and absence of signal as 0% (white), according to the Meta-profile analysis tool within Genevestigator V3.

about half of the *LecRK* genes was activated and showed much higher expression levels than the controls. In the presence of high CO_2 concentrations the reverse was observed, with many of the *LecRK* genes showing a lower expression than the controls.

Upon treatment with a variety of elicitors, rather similar patterns were obtained (Fig. 4). Expression of several *LecRK* genes from different clades was strongly induced upon treatment with six out of seven elicitors. The exception is bacterial lipopolysaccharide (LPS) which hardly induced any *LecRK* apart from *LecRK-IX.2*. The genes that were most strongly induced are *LecRK-I.8*, *LecRK-IV.2*, *LecRK-IV.3*, *LecRK-V.2*, *LecRK-V.5*, *LecRK-VI.2*, *LecRK-VI.3*, *LecRK-VII.2*, and *LecRK-IX.2*. *LecRK-VI.4* behaved remarkably; a strong repression by syringolin but a strong induction by five other elicitors.

Unlike the rather uniform expression pattern that is observed upon elicitor treatment, the patterns observed during interaction with biotic agents are much more diverse. The most strongest repression was measured upon interaction with the aphid *Myzus persicae* with 17 *LecRK* genes strongly down-regulated and none up-regulated. Interaction with another insect, the sweet potato whitefly *Bemisia tabaci,* only induced the activity of a few *LecRK* genes, and resulted in most cases in an increase in expression rather than a decrease. Virus infection influenced the expression of several *LecRK* genes, both positive and negative, and the patterns were specific for each of the two plant–virus interactions included in this analysis. No significant changes in gene expression were observed during interaction with the arbuscular mycorrhizal fungus *Gigaspora rosea* or the nematode *Meloidogyne incognita.* In addition, a compatible interaction with the powdery mildews *Erysiphe cichoracearum* or *E. orontii* did not alter *LecRK* expression substantially. By contrast, in an incompatible interaction with the non-host powdery mildew *Blumeria graminis* f. sp. *hordei* (synonym *Erysiphe graminis*) expression of several *LecRK* genes was induced or repressed. Similarly, expression was largely unaffected upon infection with the pathogenic bacterium *Pseudomonas syringae* DC3000, or its non-pathogenic type III secretion-defective mutant hrpA. However, in incompatible interactions resulting from infection with *Pseudomonas* strains secreting the effectors AvrRpt2, AvrRpm1 or AvrRps4, expression of several *LecRK* genes was found to be induced. Four of these, *LecRK-I.8*, *LecRK-I.9*, *LecRK-V.5*, and *LecRK-IX.2*, were also induced upon inoculation of *Arabidopsis* with the oomycete *Phytophthora infestans. Arabidopsis* is a non-host for *P. infestans* and this may explain why these particular *LecRK* genes were also strongly induced upon elicitor treatment. Inoculation of

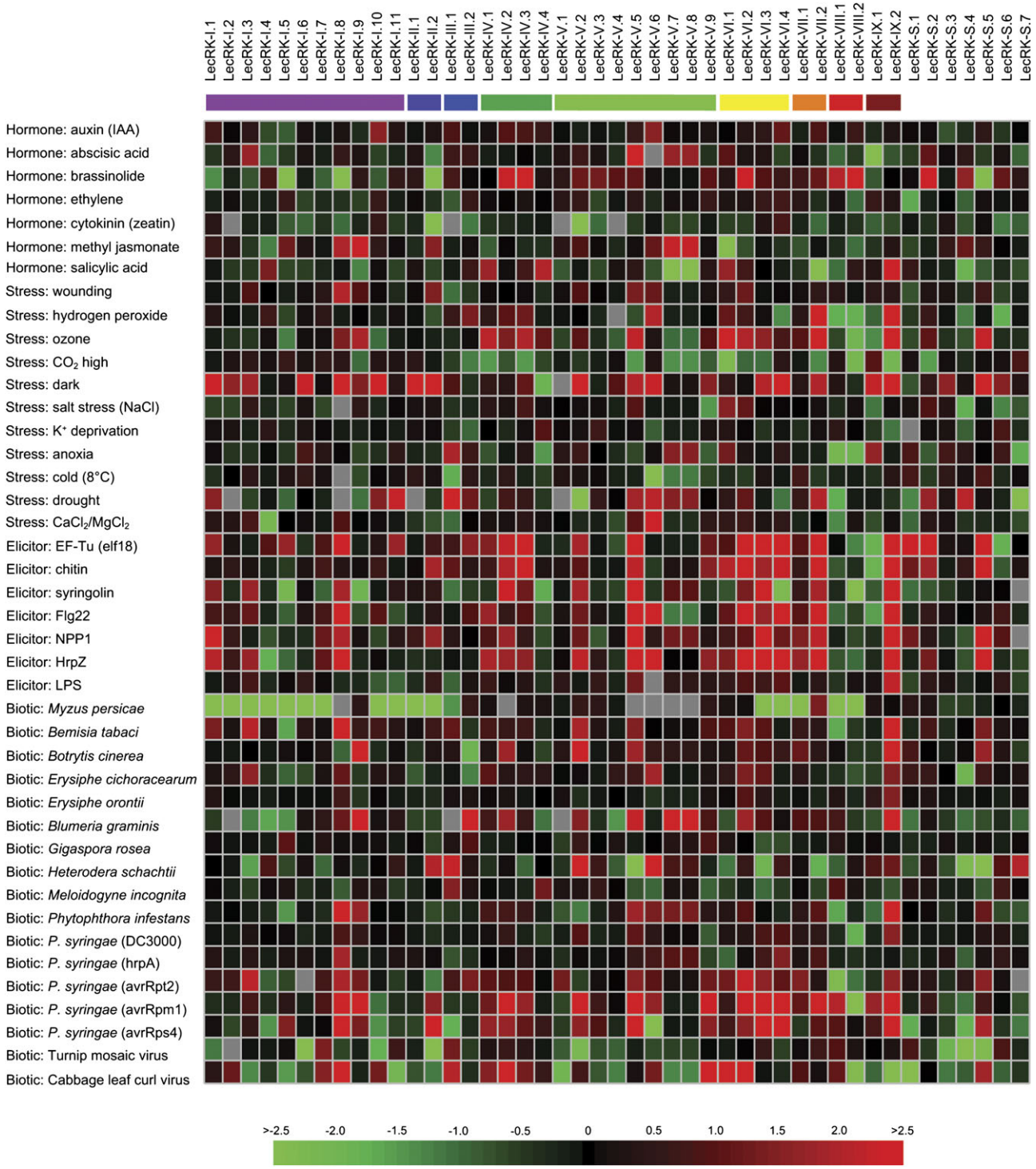

Fig. 4. Expression profiles of the *LecRK* gene family in response to hormone treatment, abiotic stress, elicitor treatment, and pathogen infection. Signal intensity $\log_2$ ratios (treatment versus control) are colour-coded (red and green for relatively higher or lower expression, respectively) as indicated in the scale bar. Black indicates no change between conditions. Grey depicts missing data points or indicates that both values for treatment and control are in the background range. The highest and lowest $\log_2$ signal value is –7.58 and 11.8, respectively.

Arabidopsis with the grey mould fungus *Botrytis cinerea* results in necrotic lesions and, as observed in Fig. 4, this was accompanied by the induction or repression of a few particular *LecRK* genes including some that are affected by other pathogens.

LecRK *genes are differentially expressed among* Arabidopsis *accessions*

To examine whether *LecRK* genes are differentially expressed among *Arabidopsis* accessions, expression data

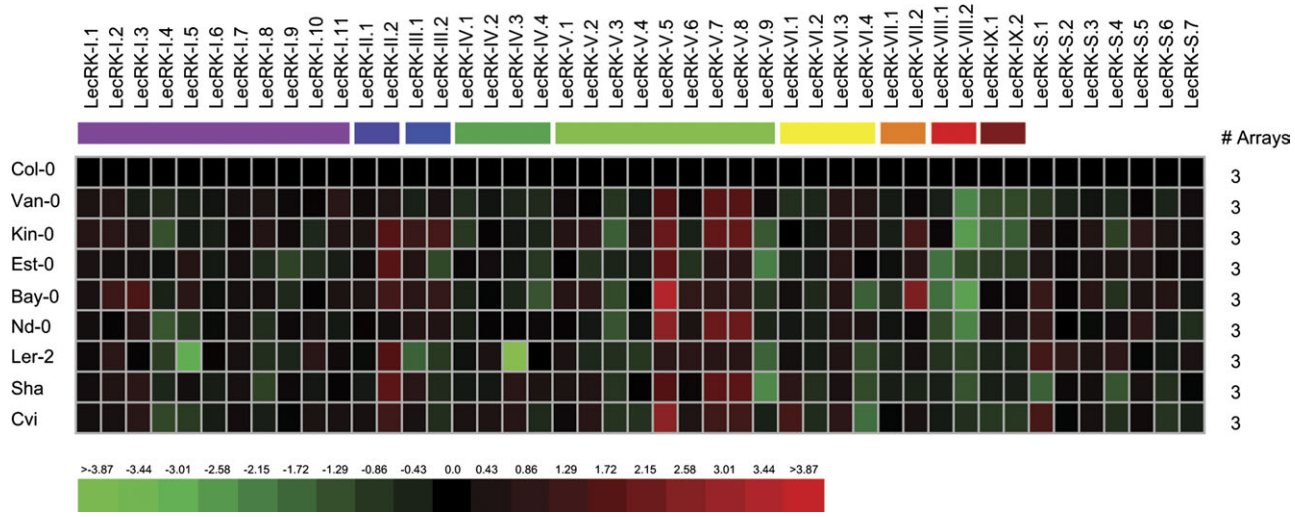

Fig. 5. Relative expression difference of *LecRK* genes between *Arabidopsis* accessions. Col-0, Columbia, USA; Van-0, Vancouver, USA; Kin-0, Kindalville, USA; Est-0, Estonia; Bay-0, Bayreuth, Germany; Nd-0, Niederzenz, Germany; L*er*-2, Landsberg, Germany; Sha, Shahdara, Tajikistan; Cvi, Cape Verde Islands. The $\log_2$ signal intensity ratios are colour-coded (red and green for relatively higher or lower expression, respectively) as indicated in the scale bar. The highest absolute value *vis-a-vis* colour code is 4.3.

were retrieved from the Bio-Array Resource (BAR) website using the *Arabidopsis* Natural Variation eFP Browser. Expression values relative to Col-0 were retrieved for each *LecRK* gene member. Only microarray data generated from experiments that made use of triplicate sampling of plant material (aerial parts of 4-d-old seedlings) were analysed and this restricted the analysis to the *Arabidopsis* accessions Col-0, Van-0, Kin-0, Est-0, Bay-0, Nd-0, L*er*-2, Sha, and Cvi. Transcript profiles were visualized with the BAR HeatMapper Plus. For most *LecRK* genes relatively similar expression levels were found among the various accessions (Fig. 5). Exceptions are *LecRK-II.2*, *LecRK-V.5*, *LecRK-IV.7*, and *LecRK-IV.8*, which seem to have a higher relative expression in all accessions when compared to Col-0. It should be noted that *LecRK-IV.7* and *LecRK-IV.8* hybridize to the same probe set resulting in equal signal intensities. The few genes with a low relative expression among the accessions are mainly in clade LecRK-VIII and LecRK-IX. Only a few *LecRK* genes were found to have a high expression value in one accession but a low expression in another accession. The most striking differences in expression values were found for *LecRK-I.5* and *LecRK-IV.3* both of which have substantially lower relative transcript levels in L*er*-2 and *LecRK-V.5* that has a higher expression in the accession Bay-0.

Conclusions

Several membrane-bound receptors in plants are shown to be functional in the perception of external signals and the subsequent triggering of downstream signal cascades to initiate adaptive responses. In this study, the focus was on the LecRK subfamily of membrane-bound receptors in *Arabidopsis*. The available gene information and expression data of the 45 members in this subfamily have been summarized and a new systematic nomenclature is proposed for the *LecRK* genes based on clades within the subfamily. Previously, Barre and co-workers (2002) conducted a similar analysis for 42 *Arabidopsis LecRK* genes, including structural alignments and molecular modelling of the encoded proteins. In that study, however, three genes that were predicted to encode LecRKs by Shiu and Bleeker (2001*b*, 2003) were not included. Here, the phylogenetic relationship within the *LecRK* gene family has been revised by constructing phylograms with more sophisticated statistical and computational methods. Three phylograms which were rooted with five outgroup RLKs and based on multi-alignments of the full-length proteins, lectin motifs, and kinase domains, respectively, showed clustering in nine different clades and seven singletons. Pairwise comparisons also resulted in the same nine clades and seven singletons and this consensus strengthens our clade division. The subdivision into clades and singletons formed the basis for the new, more systematic nomenclature and provides a helpful standard for gene identification and naming in other plant species.

The currently available expression datasets provide a basis to generate biological hypotheses as specific patterns of gene expression may give hints about their functionality. EST, MPSS, and microarray data indicate that the *LecRK* genes have variable expression patterns in different *Arabidopsis* tissues, developmental stages, in response to stimuli, and in various accessions. Several of the *LecRK*s have low transcript levels during plant development, but are induced by hormone or elicitor treatment or during interaction with pathogens. For example, a striking difference was found for the two genes in the LecRK-III clade. They are not expressed during plant development, but are expressed in response to biotic and abiotic stimuli. They do respond to different stimuli, though and, thus, their expression pattern is not clade

specific. Overall, there was no correlation between the expression pattern of a particular gene and its position in the phylogenetic tree. For all 45 *LecRK* genes, including the ones encoding unusual protein structures, expression was observed in at least one of the conditions that was analysed, hence, demonstrating that none of them represents a non-functional dead copy. So far, functional analysis of the *Arabidopsis LecRK* genes based on knock-out mutants has been limited to 5 out of the 45 family members, i.e. *LecRK-I.9*, *LecRK-IV.2*, and three members of the *LecRK-VI* clade (K Bouwmeester *et al.*, unpublished data; Wan *et al.*, 2008; Xin *et al.*, 2009). The phylogenetic relationship and the expression analysis presented in this study may help to select other candidates for more indepth studies aimed at unravelling the role of LecRKs in various biological processes, in stress responses or in interaction with pathogens.

Supplementary data

Supplementary data are available at *JXB* online.

Supplementary Table S1 contains the MPSS 17 bp tag expression data of the *Arabidopsis LecRK* genes.

Acknowledgements

We would like to acknowledge Michael Seidl and Ronnie de Jonge for insightful discussions and Dr Harold Meijer for constructive feedback on the first manuscript drafts. This research was supported by the Dutch Ministry of Agriculture, Nature and Food quality, LNV427 grant ('Parapluplan *Phytophthora*') and by an EU-BioExploit grant (FOOD-CT-2005-513959).

References

Anderson CM, Wagner TA, Perret M, He ZH, He D, Kohorn BD. 2001. WAKs: cell wall-associated kinases linking the cytoplasm to the extracellular matrix. *Plant Molecular Biology* **47,** 197–206.

André S, Siebert HC, Nishiguchi M, Tazaki K, Gabius HJ. 2005. Evidence for lectin activity of a plant receptor-like protein kinase by application of neoglycoproteins and bioinformatic algorithms. *Biochimica et Biophysica Acta* **1725,** 222–232.

Asselbergh B, De Vleesschauwer D, Höfte M. 2008. Global switches and fine-tuning: ABA modulates plant pathogen defence. *Molecular Plant–Microbe Interactions* **21,** 709–719.

Barre A, Hervé C, Lescure B, Rougé P. 2002. Lectin receptor kinases in plants. *Critical Reviews in Plant Sciences* **21,** 379–399.

Bouwmeester K, Klamer S, Gouget A, Haget N, Canut H, Govers F. 2008. Lectin receptor kinase 79, a putative target of the *Phytophthora infestans* effector IPI-O. In: Lorito M, Woo SL, Scala F, eds. *Biology of plant–microbe interactions*. St Paul, MN, USA: International Society for Molecular Plant–Microbe Interactions.

Cambi A, Koopman M, Figdor CG. 2005. How C-type lectins detect pathogens. *Cellular Microbiology* **7,** 481–488.

Chen X, Shang J, Chen D, *et al*. 2006. A B-lectin receptor kinase gene conferring rice blast resistance. *The Plant Journal* **46,** 794–804.

Cordeiro MCR, Piqueras R, de Oliveira DE, Castresana C. 1998. Characterization of early induced genes in *Arabidopsis thaliana* responding to bacterial inoculation: identification of centrin and of a novel protein with two regions related to kinase domains. *FEBS Letters* **434,** 387–393.

Decreux A, Messiaen J. 2005. Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. *Plant and Cell Physiology* **46,** 268–278.

Gómez-Gómez L, Boller T. 2000. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell* **5,** 1003–1011.

Gouget A, Senchou V, Govers F, Sanson A, Barre A, Rougé P, Pont-Lezica R, Canut H. 2006. Lectin receptor kinases participate in protein–protein interactions to mediate plasma membrane-cell wall adhesions in Arabidopsis. *Plant Physiology* **140,** 81–90.

He XJ, Zhang ZG, Yan DQ, Zhang JS, Chen SY. 2004. A salt-responsive receptor-like kinase gene regulated by the ethylene signalling pathway encodes a plasma membrane serine/threonine kinase. *Theoretical and Applied Genetics* **109,** 377–383.

Hématy K, Sado PE, Van Tuinen A, Rochange S, Desnos T, Balzergue S, Pelletier S, Renou JP, Höfte H. 2007. A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Current Biology* **17,** 922–931.

Hervé C, Dabos P, Galaud JR, Rougé P, Lescure B. 1996. Characterization of an *Arabidopsis thaliana* gene that defines a new class of putative plant receptor kinases with an extracellular lectin-like domain. *Journal of Molecular Biology* **258,** 778–788.

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. 2008. Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics* doi: 10.1155/2008/420747.

Humphrey TV, Bonetta DT, Goring DR. 2007. Sentinels at the wall: cell wall receptors and sensors. *New Phytologist* **176,** 7–21.

Kanzaki H, Saitoh H, Takahashi Y, Berberich T, Ito A, Kamoun S, Terauchi R. 2008. NbLRK1, a lectin-like receptor kinase protein of *Nicotiana benthamiana*, interacts with *Phytophthora infestans* INF1 elicitin and mediates INF1-induced cell death. *Planta* **228,** 977–987.

Kim Y-T, Oh J, Kim K-H, Uhm J-Y, Lee B-M. 2009. Isolation and characterization of *NgRLK1*, a receptor-like kinase of *Nicotiana glutinosa* that interacts with the elicitin of *Phytophthora capsici*. *Molecular Biology Reports*.

Kinoshita T, Cano-Delgado A, Seto H, Hiranuma S, Fujioka S, Yoshida S, Chory J. 2005. Binding of brassinosteroids to the extracellular domain of plant receptor kinase BRI1. *Nature* **433,** 167–171.

Kohorn BD, Kobayashi M, Johansen S, Friedman HP, Fischer A, Byers N. 2006*a*. Wall-associated kinase 1 (WAK1) is crosslinked in endomembranes, and transport to the cell surface requires correct cell-wall synthesis. *Journal of Cell Science* **119,** 2282–2290.

Kohorn BD, Kobayashi M, Johansen S, Riese J, Huang LF, Koch K, Fu S, Dotson A, Byers N. 2006*b*. An Arabidopsis cell wall-associated kinase required for invertase activity and cell growth. *The Plant Journal* **46,** 307–316.

Kusaba M, Dwyer K, Hendershot J, Vrebalov J, Nasrallah JB, Nasrallah ME. 2001. Self-incompatibility in the genus *Arabidopsis*: characterization of the *S* locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. *The Plant Cell* **13,** 627–643.

Lally D, Ingmire P, Tong HY, He ZH. 2001. Antisense expression of a cell wall-associated protein kinase, WAK4, inhibits cell elongation and alters morphology. *The Plant Cell* **13,** 1317–1331.

Larkin MA, Blackshields G, Brown NP, *et al*. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948.

Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ. 2005. Elucidation of the small RNA component of the transcriptome. *Science* **309,** 1567–1569.

Mondragon-Palomino M, Gaut BS. 2005. Gene conversion and the evolution of three leucine-rich repeat gene families in *Arabidopsis thaliana*. *Molecular Biology and Evolution* **22,** 2444–2456.

Naithani S, Chookajorn T, Ripoll DR, Nasrallah JB. 2007. Structural modules for receptor dimerization in the *S*-locus receptor kinase extracellular domain. *Proceedings of the National Academy of Sciences, USA* **104,** 12211–12216.

Nakhamchik A, Zhao Z, Provart NJ, Shiu SH, Keatley SK, Cameron RK, Goring DR. 2004. A comprehensive expression analysis of the Arabidopsis proline-rich extensin-like receptor kinase gene family using bioinformatic and experimental approaches. *Plant and Cell Physiology* **45,** 1875–1881.

Navarro-Gochicoa MT, Camut S, Timmers ACJ, Niebel A, Hervé C, Boutet E, Bono JJ, Imberty A, Cullimore JV. 2003. Characterization of four lectin-like receptor kinases expressed in roots of *Medicago truncatula*. Structure, location, regulation of expression, and potential role in the symbiosis with *Sinorhizobium meliloti*. *Plant Physiology* **133,** 1893–1910.

Ogawa M, Shinohara H, Sakagami Y, Matsubayash Y. 2008. Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. *Science* **319,** 294.

Park AR, Cho SK, Yun UJ, Jin MY, Lee SH, Sachetto-Martins G, Park OK. 2001. Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. *Journal of Biological Chemistry* **276,** 26688–26693.

Riou C, Hervé C, Pacquit V, Dabos P, Lescure B. 2002. Expression of an Arabidopsis lectin kinase receptor gene, *lecRK-a1*, is induced during senescence, wounding and in response to oligogalacturonic acids. *Plant Physiology and Biochemistry* **40,** 431–438.

Schulze-Muth P, Irmler S, Schröder G, Schröder J. 1996. Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*): cDNA, gene, intramolecular autophosphorylation, and identification of a threonine important for auto- and substrate phosphorylation. *Journal of Biological Chemistry* **271,** 26684–26689.

Senchou V, Weide R, Carrasco A, Bouyssou H, Pont-Lezica R, Govers F, Canut H. 2004. High affinity recognition of a *Phytophthora* protein by Arabidopsis via an RGD motif. *Cellular and Molecular Life Sciences* **61,** 502–509.

Sherman-Broyles S, Boggs N, Farkas A, Liu P, Vrebalov J, Nasrallah ME, Nasrallah JB. 2007. *S* locus genes and the evolution of self-fertility in *Arabidopsis thaliana*. *The Plant Cell* **19,** 94–106.

Shiu SH, Bleecker AB. 2001*a*. Plant receptor-like kinase gene family: diversity, function, and signalling. *Science's STKE* **113,** re22.

Shiu SH, Bleecker AB. 2001*b*. Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proceedings of the National Academy of Sciences, USA* **98,** 10763–10768.

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

Swarup R, Dumas C, Cock JM. 1996. A new class of receptor-like protein kinase gene from *Arabidopsis thaliana* possessing a domain similarity to plant lectin genes (PGR96-022). *Plant Physiology* **111,** 347.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24,** 1596–1599.

Van Damme EJM, Nakamura-Tsuruta S, Smith DF, *et al*. 2007. Phylogenetic and specificity studies of two-domain GNA-related lectins: generation of multispecificity through domain duplication and divergent evolution. *Biochemical Journal* **404,** 51–61.

Verica JA, He ZH. 2002. The cell wall-associated kinase (WAK) and WAK-like kinase gene family. *Plant Physiology* **129,** 455–459.

Wan J, Patel A, Mathieu M, Kim SY, Xu D, Stacey G. 2008. A lectin receptor-like kinase is required for pollen development in Arabidopsis. *Plant Molecular Biology* **67,** 469–682.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. An 'electronic fluorescent pictograph' browser for exploring and analysing large-scale biological data sets. *PLoS ONE* **2,** e718.

Xin Z, Wang A, Yang G, Gao P, Zheng ZL. 2009. The Arabidopsis A4 subfamily of lectin receptor kinases negatively regulates abscisic acid response in seed germination. *Plant Physiology* **149,** 434–444.

Yamaoka K, Saharinen P, Pesu M, Holt III VET, Silvennoinen O, O'Shea JJ. 2004. The Janus kinases (Jaks). *Genome Biology* **5,** 253.

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JDG, Boller T, Felix G. 2006. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell* **125,** 749–760.