Molecular Plant Pathology

MOLECULAR PLANT PATHOLOGY (2013) **14**(6), 602–609 DOI: 10.1111/mpp.12028

Short communication

Novel receptor-like kinases in cacao contain PR-1 extracellular domains

PAULO JOSÉ PEREIRA LIMA TEIXEIRA[1], GUSTAVO GILSON LACERDA COSTA[1], GABRIEL LORENCINI FIORIN[1], GONÇALO AMARANTE GUIMARÃES PEREIRA[1] AND JORGE MAURÍCIO COSTA MONDEGO[2,*]

[1]*Laboratório de Genômica e Expressão, Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas (Unicamp), CP 6109, Campinas, SP, 13083-970, Brazil*

[2]*Centro de Pesquisa e Desenvolvimento em Recursos Genéticos Vegetais, Instituto Agronômico de Campinas (IAC), Campinas, SP, 13075-630, Brazil*

SUMMARY

Members of the pathogenesis-related protein 1 (PR-1) family are well-known markers of plant defence responses, forming part of the arsenal of the secreted proteins produced on pathogen recognition. Here, we report the identification of two cacao (*Theobroma cacao* L.) PR-1s that are fused to transmembrane regions and serine/threonine kinase domains, in a manner characteristic of receptor-like kinases (RLKs). These proteins (TcPR-1f and TcPR-1g) were named PR-1 receptor kinases (PR-1RKs). Phylogenetic analysis of RLKs and PR-1 proteins from cacao indicated that PR-1RKs originated from a fusion between sequences encoding PR-1 and the kinase domain of a LecRLK (Lectin Receptor-Like Kinase). Retrotransposition marks surround *TcPR-1f*, suggesting that retrotransposition was involved in the origin of PR-1RKs. Genes with a similar domain architecture to cacao PR-1RKs were found in rice (*Oryza sativa*), barrel medic (*Medicago truncatula*) and a nonphototrophic bacterium (*Herpetosiphon aurantiacus*). However, their kinase domains differed from those found in LecRLKs, indicating the occurrence of convergent evolution. *TcPR-1g* expression was up-regulated in the biotrophic stage of witches' broom disease, suggesting a role for PR-1RKs during cacao defence responses. We hypothesize that PR-1RKs transduce a defence signal by interacting with a PR-1 ligand.

Pathogenesis-related (PR) proteins are produced by plants during pathogen infection. Members of the PR-1 family are well-known markers of plant defence responses (Van Loon *et al*., 2006) and some of these have been found to inhibit the development of fungi and oomycetes *in vitro* (Niderman *et al*., 1995; Rauscher *et al*., 1999). Although these proteins are widely recognized as components of the plant defence system, their function has not been well defined. PR-1 proteins are part of the SCP/TAPS (Sperm-Coating Protein/Tpx-1-Ag5-PR-1-Sc7) superfamily, which consists of proteins from a wide phylogenetic spectrum involved in defence responses, pathogenesis and reproduction (Cantacessi *et al*., 2009; Gibbs *et al*., 2008). In addition to a conserved SCP/TAPS domain, many proteins from the SCP/TAPS superfamily contain C-terminal extensions that have functional domains (Gibbs *et al*., 2008). Curiously, the plant PR-1 proteins experimentally characterized to date contain only the SCP/TAPS domain, and lack any C-terminal domain extension.

Cacao (*Theobroma cacao* L.) is one of the most economically important perennial crops in the world, producing the main feedstock for the chocolate industry (Afoakwa *et al*., 2008). However, cacao production is severely hindered by diseases caused by fungi and oomycetes (Evans, 2007). Hence, the study of proteins associated with plant immunity can provide clues on the development of strategies to control pathogen infection in cacao plantations. Based on evidence that *PR-1* genes are related to defence responses in plants (Van Loon *et al*., 2006), we searched for these genes in the cacao genome. First, we used the sequence of the tomato PR-1 protein P14 (Fernández *et al*., 1997; GenBank P04284) as bait in a TBLASTN search of the Cacao Genome Database (http://www.cacaogenomedb.org), a consortium from MARS/US Department of Agriculture-Agricultural Research Service that sequenced the Matina 1–6 genotype. Matina 1–6 is representative of the Forastero genetic background most commonly found in cacao-producing countries. Thirteen genes with significant sequence similarity (BLASTP positives > 55%; *E*-value < 1e-25) to P14 were identified in the Matina 1–6 cacao genome (Table 1). We named these genes *TcPR-1a* to *TcPR-1m*. Consistent with BLAST analyses, a search for members of the SCP/TAPS family (InterPro IPR001283) using the InterProScan server (Quevillon *et al*., 2005; http://www.ebi.ac.uk/Tools/pfa/iprscan) also revealed the existence of 13 protein models in cacao. The InterPro analysis also revealed that the 13 PR-1 proteins from *T. cacao* have predicted hydrophobic signal peptides, indicating that these proteins are potentially directed to the secretory pathway. In addition, all of these proteins contain complete SCP/TAPS domains. Strikingly,

**Correspondence*: Email: jmcmondego@gmail.com; jmcmondego@iac.sp.gov.br

two TcPR-1s (TcPR-1f and TcPR-1g) contain extensions in the C-terminal portion that are similar to a serine/threonine kinase (S-TKc; InterPro IPR000719; Fig. 1a) domain. Both TcPR-1f and TcPR-1g contain a stretch of 23 hydrophobic amino acids that forms a membrane-spanning helix between the SCP/TAPS and S-TKc domains, strongly suggesting that these proteins are anchored in a cell membrane (Fig. 1a). To determine the cellular localization of these proteins, we performed an *in silico* analysis using the PSORT program (Nakai and Horton, 1999). As expected, these two proteins localized to the plasma membrane. Such protein architecture (i.e. an extracellular region, a hydrophobic stretch and a kinase domain) is characteristic of receptor-like kinases (RLKs).

RLKs are a large group of transmembrane receptors that perceive extracellular signals, which are transduced via phosphorylation activation and lead to intracellular responses (Walker, 1994). RLKs are one of the largest gene families in plants. For instance, the Arabidopsis genome encodes approximately 600 RLKs, whereas more than 1100 RLK genes are present in the rice genome (Dardick *et al*., 2007; Gish and Clark, 2011; Morillo and Tax, 2006; Shiu and Bleecker, 2001, 2003; Shiu *et al*., 2004). RLKs have been organized into different structural classes, based on the identity of their extracellular domain and on phylogenetic relationships of their kinase domain (Shiu and Bleecker, 2001). These receptors underwent tremendous expansion and diversification, mostly in their extracellular domains, indicating the necessity of variation in the extracellular binding domains for the recognition of a vast range of ligands (Lehti-Shiu *et al*., 2009; Shiu and Bleecker, 2003), which can be brassinosteroids (Wang *et al*., 2001), polysaccharides such as chitin (Miya *et al*., 2007) or peptides such as CLAVATA 3 (Trotochaud *et al*., 2000), flagellin (Gomez-Gomez and Boller, 2000; Zipfel *et al*., 2004) and the bacterial elongation factor EF-Tu (Zipfel *et al*., 2006).

Table 1 Characteristics of the 13 pathogenesis-related protein 1 (PR-1) gene models found in *Theobroma cacao*.

TcPR-1	Gene ID*	Chromosome	Putative protein size (amino acids)	Putative coding sequence size (bp)
TcPR-1a	CGD0027643	7	159	477
TcPR-1b	CGD0027642	7	156	468
TcPR-1c	CGD0027635	7	162	486
TcPR-1d	CGD0027628	7	164	492
TcPR-1e	CGD0027640	7	165	495
TcPR-1f	CGD0008870	2	619	1857
TcPR-1g	CGD0006833	10	582	1746
TcPR-1h	CGD0021343	5	173	519
TcPR-1i	CGD0021746	5	236	708
TcPR-1j	CGD0031974	9	185	555
TcPR-1k	CGD0013072	†	195	585
TcPR-1l	CGD0000407	1	177	531
TcPR-1m	CGD0027644	7	159	477

*Gene IDs according to the Cacao Genome Database (v0.9 – http://www.cacaogenomedb.org).

†This gene model was mapped in a *T. cacao* supercontig (super_217) that was not assembled in any cacao chromosome.

Many RLKs are involved in developmental processes in plants, such as floral organ abscission (HAESA, Jinn *et al*., 2000) and reproduction (FERONIA; Escobar-Restrepo *et al*., 2007). Other RLKs are part of the defensive arsenal against pathogens (Afzal *et al*., 2008; Greeff *et al*., 2012; Torii, 2004). RLKs involved in plant immunity are so-called pattern-recognition receptors (PRRs). They detect microbe- or pathogen-associated molecular patterns (MAMPs/PAMPs) and, on binding of their cognate ligands, set off defence responses (Boller and Felix, 2009; Greeff *et al*., 2012; Monaghan and Zipfel, 2012). Some reports have shown that the loss of function of RLKs enhances susceptibility to pathogens (Chen *et al*., 2006; Wan *et al*., 2008; Zipfel *et al*., 2004), attesting to their role as plant immune receptors. However, ligands have been identified for only a few RLK-PRRs: FLS2 (Zipfel *et al*., 2004), Xa21 (Lee *et al*., 2009), EFR (Zipfel *et al*., 2006) and CERK1 (Miya *et al*., 2007). Molecules such as chitin (Kaku *et al*., 2006) and peptidoglycan (Liu *et al*., 2012; Willmann *et al*., 2011) have been shown to interact with receptor-like proteins (RLPs), which are transmembrane receptors that lack a kinase domain (Wang *et al*., 2008).

To confirm the protein models of cacao PR-1 RLKs (hereafter referred to as PR-1RK), we designed oligonucleotides based on the sequences of the N- and C-terminals of TcPR-1f and TcPR-1g and used as templates genomic DNA and cDNA isolated from plantlets of cv. Comum, an Amazonic variety of the Forastero group of cacao (Vello and Garcia, 1971), which is very similar to the Matina 1–6 cultivar (Soria, 1996). A comparison of the genomic and cDNA amplifications revealed that *TcPR-1RKs* contain no introns spanning their coding sequences.

We examined the TcPR-1f and TcPR-1g protein sequences for features of S-TKcs. The cytoplasmic region of both proteins contains the 11 kinase subdomains typical of S-TKcs described by Walker (1994), including the subdomain I (ATP-binding site; consensus GXGXXG), and subdomains VIb (consensus DxKxxN) and VIII (consensus GTxxYxAPE), which potentially discriminate S-TKcs from tyrosine kinases (Fig. 1a). Thus, both proteins are candidate S-TKcs. However, some plant RLKs that were predicted to be S-TKcs (e.g. BAK1 and BRI1) have been shown to be dual-specificity kinases (Oh *et al*., 2009, 2010). In this context, functional assays are necessary to confirm the specificity of the cacao PR-1RK kinase domains.

In addition, both PR-1RKs can be classified as RD kinases, because they contain an arginine (R) residue preceding the core catalytic aspartate (D) residue in subdomain VIb (Johnson *et al*., 1996; Fig. 1a). Although most kinases involved in microbial detection are non-RD kinases (lacking the arginine residue before the catalytic aspartate), RD kinases, such as ERECTA (Godiard *et al*., 2003), WAKL22 (Diener and Ausubel, 2005) and, especially, BAK1 (Chinchilla *et al*., 2007; Heese *et al*., 2007; Roux *et al*., 2011), have been found to contribute to pathogen resistance.

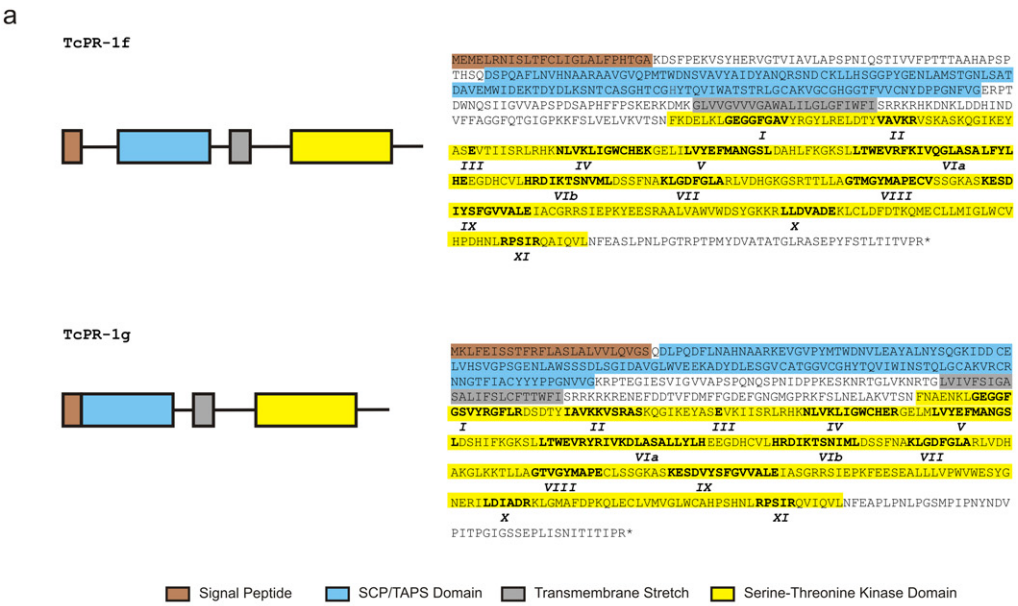

a
TcPR-1f
MEMELRNISLTFCLIGLALFPHTGAKDSFPEKVSYHERVGTVIAVLAPSPNIQSTIVVFPTTTAAHAPSP
THSQDSPQAFLNVHNAARAAVGVQPMTWDNSVAVYAIDYANQRSNDCKLLHSGGPYGENLAMSTGNLSAT
DAVEMWIDEKTDYDLKSNTCASGHTCGHYTQVIWATSTRLGCAKVGCGHGGTFVVCNYDPPGNFVGERPT
DWNQSIIGVVAPSPDSAPHFFPSKERKDMKGLVVGVVVGAWALILGLGFIWFISRRKRHKDNKLDDHIND
VFFAGGFQTGIGPKKFSLVELVKVTSNFKDELKLGEGGFGAVYRGYLRELDTYVAVKRVSKASKQGIKEY
I II
ASEVTIISRLRHKNLVKLIGWCHEKGELILVYEFMANGSLDAHLFKGKSLLTWEVRFKIVQGLASALFYL
III IV V VIa
HEEGDHCVLHRDIKTSNVMLDSSFNAKLGDFGLARLVDHGKGSRTTLLAGTMGYMAPECVSSGKASKESD
VIb VII VIII
IYSFGVVALEIACGRRSIEPKYEESRAALVAWVWDSYGKKRLLDVADEKLCLDFDTKQMECLLMIGLWCV
IX X
HPDHNLRPSIRQAIQVLNFEASLPNLPGTRPTPMYDVATATGLRASEPYFSTLTITVPR*
XI
TcPR-1g
MKLFEISSTFRFLASLALVVLQVGSQDLPQDFLNAHNAARKEVGVPYMTWDNVLEAYALNYSQGKIDDCE
LVHSVGPSGENLAWSSSDLSGIDAVGLWVEEKADYDLESGVCATGGVCGHYTQVIWINSTQLGCAKVRCR
NNGTFIACYYYPPGNVVGKRPTEGIESVIGVVAPSPQNQSPNIDPPKESKNRTGLVIVFSIGA
SALIFSLCFTTWFISRRKRKRENEFDDTVFDMFFGDEFGNGMGPRKFSLNELAKVTSNFNAENKLGEGGF
GSVYRGFLRDSDTYIAVKKVSRASKQGIKEYASEVKIISRLRHKNLVKLIGWCHERGELMLVYEFMANGS
I II III IV V
LDSHIFKGKSLLTWEVRYRIVKDLASALLYLHEEGDHCVLHRDIKTSNIMLDSSFNAKLGDFGLARLVDH
VIa VIb VII
AKGLKKTLLAGTVGYMAPECLSSGKASKESDVYSFGVVALEIASGRRSIEPKFEESEALLLVPWVWESYG
VIII IX
NERILDIADRKLGMAFDPKQLECLVMVGLWCAHPSHNLRPSIRQVIQVLNFEAPLPNLPGSMPIPNYNDV
X XI
PITPGIGSSEPLISNITITIPR*
Signal Peptide
SCP/TAPS Domain
Transmembrane Stretch
Serine-Threonine Kinase Domain
b
LRR-RLK (IPR001611)
PERK (Pr01217)*
LysM (IPR018392)
WAK (IPR025287)
CRINKLY4-Like (IPR009091)
CrRLK1/FERONIA-Like (IPR024788)
C-type Lectin (IPR001304)
Legume Lectin (IPR001220)
S-Domain SRK (IPR001480)
CRR/DUF26 (IPR002902)
GmSARK-Like (IPR021720)
CHRK (IPR017853)
CGD0016565
CGD0034511
CGD0034522
CGD0034519
CGD0021863
CGD0021868
CGD0006397
CGD0021860
CGD0021855
CGD0009023
CGD0022080
CGD0022083
CGD0026169
CGD0026177
CGD0013434
CGD0023896
CGD0023898
CGD0023902
CGD0009091
CGD0018499
CGD0029040
CGD0021644
CGD0027497
CGD0008886
CGD0008900
CGD0008870 (TcPR-1f)
CGD0006833 (TcPR-1g)
CGD0008863
CGD0008890
CGD0027179
CGD0027186
CGD0000725
CGD0000730
CGD0000731
CGD0000733

Fig. 1 Pathogenesis-related protein 1 receptor-like kinase (PR-1RK) domain signature and cacao receptor-like kinase (RLK) phylogenetic tree. (a) Domain organization and protein kinase signatures of *Theobroma cacao* (Tc)PR-1RK proteins. The signal peptide (SP) is indicated in brown, the hydrophobic transmembrane region (TM) in grey and SCP/TAPS (Sperm-Coating Protein/Tpx-1-Ag5-PR-1-Sc7) (PR-1-like) and serine/threonine kinase domains in blue and yellow, respectively. The 11 (*I-XI*) typical motifs of protein kinases are indicated below the amino acids depicted in bold. (b) Phylogenetic analysis of RLKs from cacao. The branches of the phylogenetic tree are coloured according to the identity of the extracellular domains of RLKs as predicted by InterProScan software (left legend). The right panel depicts an expanded view of the legume lectin (Lectin Receptor-Like Kinases, LecRLKs) clade. The yellow box depicts PR-1RK proteins clustered together with LecRLKs. The tree was constructed as a consensus of 10 000 bootstrap replicates, using neighbour joining and distance *p* parameters. Bootstrap values are shown in the LecRLK clade. *PR01217 is a PRINTs database domain corresponding to the former InterPro domain IPR002965, which was removed from the database.

To investigate the origin of cacao PR-1RK proteins, we searched for RLKs in the Cacao Genome Database (http://www.cacaogenomedb.org). For this, we used the InterProScan server to identify cacao proteins containing a kinase domain (IPR000719) and one of the InterPro domains commonly found in the extracellular region of plant RLKs (Fig. 1b). For example, CHRK (chitinase receptor kinases) receptors contain kinase (IPR000719) and glycoside hydrolase (IPR017853) domains. Based on this analysis, we found that cacao has at least 480 putative RLKs. We then extracted the kinase domain sequences of these proteins and aligned them using Clustal Omega (Sievers *et al*., 2011). The weighing matrix used was BLOSUM62 and the alignments generated were manually adjusted following the subdomain signatures of eukaryotic kinases. Next, we used MEGA5 (Tamura *et al*., 2011) to generate a phylogenetic tree by means of the neighbour-joining method with 10 000 bootstrap replicates (Fig. 1b). Distance *p* and complete deletion parameters were applied. To classify the cacao sequences into RLK subfamilies, each tree leaf node was annotated according to previous studies (Lehti-Shiu *et al*., 2009; Shiu and Bleecker, 2003). Curiously, no cacao protein was similar to the RLK containing a thaumatin extracellular domain. TcPR-1f and TcPR-1g fit into the legume lectin (Lectin Receptor-Like Kinase, LecRLK) clade (Fig. 1b), suggesting that PR-1RKs arose from these RLKs.

We also inspected the position of PR-1RK genes in the genome of cacao Matina 1–6. *TcPR-1f* is located on chromosome 2. By examining the genomic location of this gene, we identified four *LecRLKs* in its vicinity, three downstream and one upstream of the gene (Fig. 2a). Interestingly, the LecRLK genes forming this cluster are phylogenetically close to PR-1RKs (Fig. 1b). These LecRLKs share 65%–70% sequence identity with both TcPR-1f and TcPR1-g (Fig. 2a). The genomic proximity of *LecRLKs* with a *PR-1RK* reinforces the notion that *PR-1RKs* may have originated from these *RLKs*. *TcPR-1g* is located on chromosome 10. In contrast with *TcPR-1f*, no *RLKs* were found close to this gene, which is surrounded by genes encoding an *N*-acetylglucosaminidase and a LOB domain-containing protein.

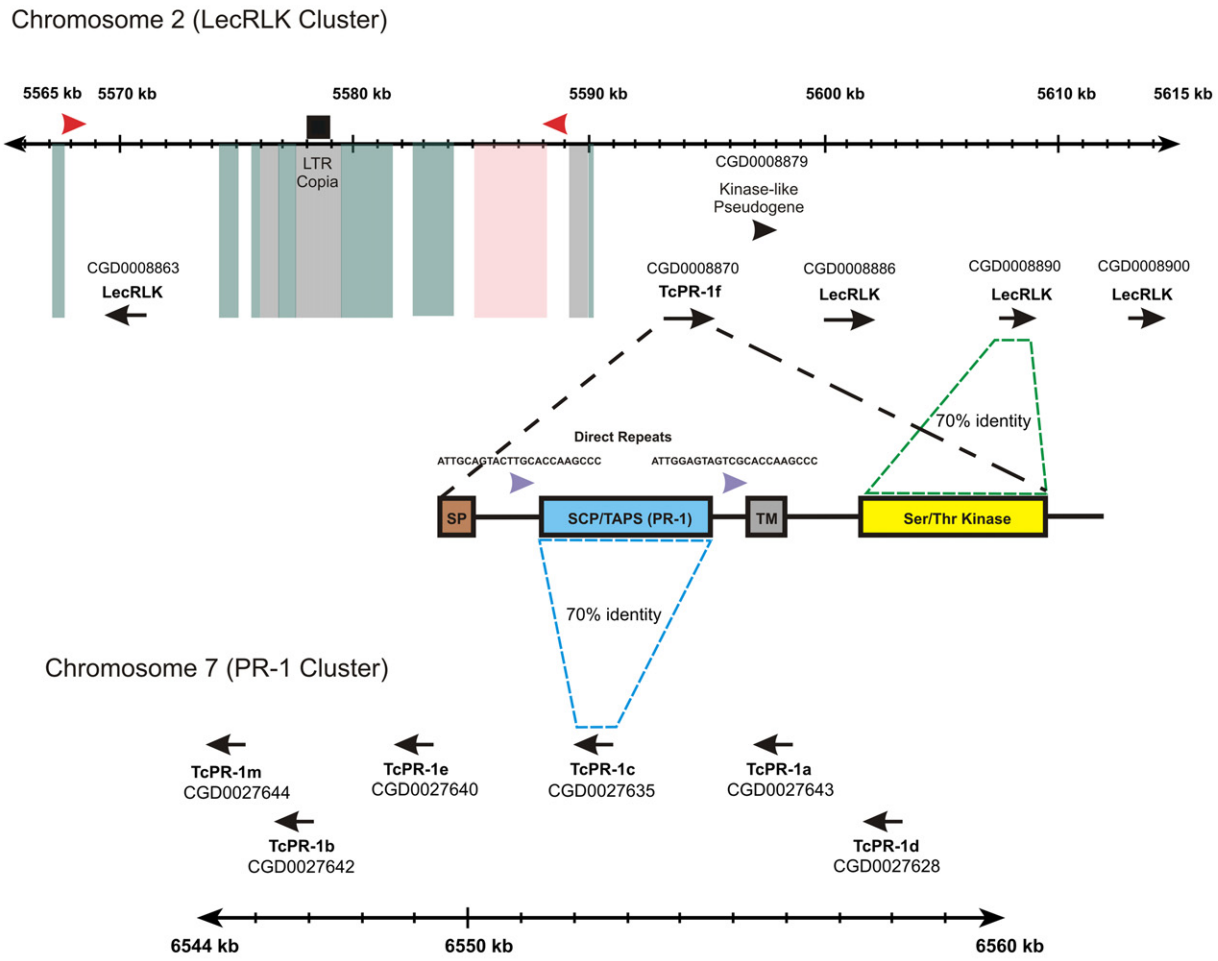
We also verified the location of the 11 secreted PR-1 proteins in the cacao genome. Five of these occur in a cluster on chromosome 7 (Table 1; Fig. 2a). An alignment of the SCP/TAPS domain of TcPR-1 proteins was used to construct a bootstrap consensus tree inferred from 5000 replicates. This analysis indicated that TcPR-1f and TcPR-1g are part of a clade containing proteins of the PR-1 cluster on chromosome 7 (Fig. 2b). The PR-1 domain of TcPR-1f shares approximately 60%–70% sequence identity with PR-1 genes positioned in tandem on chromosome 7. This result suggests that genes similar to those forming the TcPR-1 cluster fused with a kinase domain of a LecRLK, giving rise to PR-1RKs.

The formation of new chimeric proteins has been related to molecular mechanisms, such as exon shuffling, gene duplication, retrotransposition or combinations of these (Jones *et al*., 2005; Long *et al*., 2003; Nisole *et al*., 2004; Wang *et al*., 2006). Therefore, we searched for retrotransposition marks in both LecRLK and PR-1 clusters, using RepeatMasker (A.F. Smit *et al*., unpublished data; http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) and REPbase tools (Jurka, 1998). We identified a region of 770 bp that was similar to the LTR Copia-like retrotransposon in the LecRLK cluster (Fig. 2a). In addition, we detected a pair of inverted repeats that flanks the most upstream *LecRLK* in the cluster and a tandem repeat region of approximately 3800 bp (Fig. 2a). Curiously, the *TcPR-1f* PR-1 domain is flanked by direct repeats, which have been associated with transposon insertion sites (Warren *et al*., 1997). We did not find evidence of transposition events in the PR-1 cluster or close to *TcPR-1g*. The presence of retrotransposition marks and repetitive elements around *TcPR-1f* suggests that transposition mechanisms could be involved in the origin of PR-1RKs.

To evaluate if other plant species contain proteins with a domain architecture similar to the cacao PR-1RKs, we conducted a search in the Entrez Protein Database using CDART (Geer *et al*., 2002). Two proteins from the *Oryza sativa* Japonica group (EEE56976; NP_001172960) and one from the Indica group (EAY85816) have a domain structure similar to that of PR-1RKs, but contain two SCP/TAPS domains instead of one. Curiously, one protein from the legume barrel medic (*Medicago truncatula*; XP_003608315) and one from the nonphototrophic predatory bacterium *Herpetosiphon aurantiacus* (YP_001547555) contain an inverse arrangement of PR-1RK domains, having an N-terminal kinase domain followed by a C-terminal SCP/TAPS domain.

To establish whether fusions of rice and *M. truncatula* PR-1RKs also occurred between PR-1-like proteins and LecRLKs, we performed BLAST analyses using the kinase domains of these proteins. All hits from rice and *M. truncatula* PR-1RKs were similar to the cysteine-rich repeat (CRR)-RLK and leucine-rich repeat (LRR)-RLK subfamilies, respectively. The presence of PR-1RKs in rice with kinase domains similar to CRR-RLKs has been reported previously

a



b

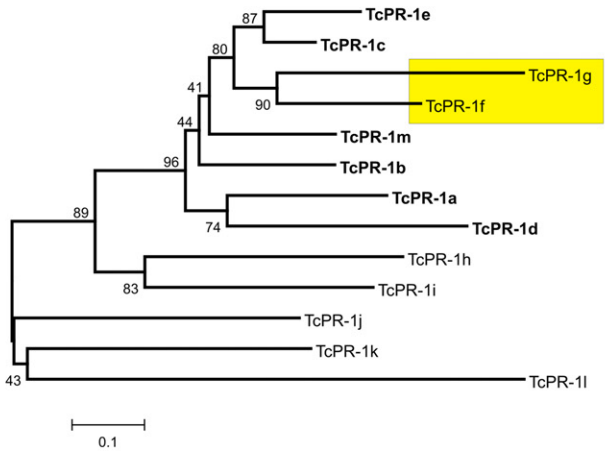


Fig. 2 Lectin Receptor-Like Kinase (*LecRLK*) and *Theobroma cacao* pathogenesis-related protein 1 (*TcPR-1*) clusters in the cacao genome and the proposed origin of pathogenesis-related protein 1 receptor-like kinases (PR-1RKs). (a) *TcPR-1f* is located in a cluster of LecRLKs on chromosome 2 of cacao. The *TcPR-1f* kinase domain shares a high level of sequence identity with kinase domains of the LecRLKs located in this cluster (i.e. CGD000890; green broken line). However, the PR-1 domain of *TcPR-1f* shares a high level of sequence identity with genes present in a PR-1 cluster on *T. cacao* chromosome 7 (i.e. TcPR-1c; blue broken line). The LecRLK cluster has signatures commonly associated with transposition events: black box, long terminal repeat (LTR) of Copia-like retrotransposon; dark grey shading, repetitive regions with at least 300 copies in the cacao genome; dark green shading, sequencing gaps; red arrowheads, inverted repeats; rose shading, tandem repeat region; lilac arrowheads, direct repeats flanking the SCP/TAPS (Sperm-Coating Protein/Tpx-1-Ag5-PR-1-Sc7) (PR-1) domain of *TcPR-1f*. (b) Phylogenetic analysis of cacao PR-1 proteins. Members that form the PR-1 cluster are depicted in bold. Notice the location of PR-1RKs (TcPR-1f and TcPR-1g; yellow box) in the clade of PR-1 cluster genes. The tree was constructed as a consensus of 5000 bootstrap replicates, using neighbour joining and distance *p* parameters. Bootstrap values are shown in the tree.

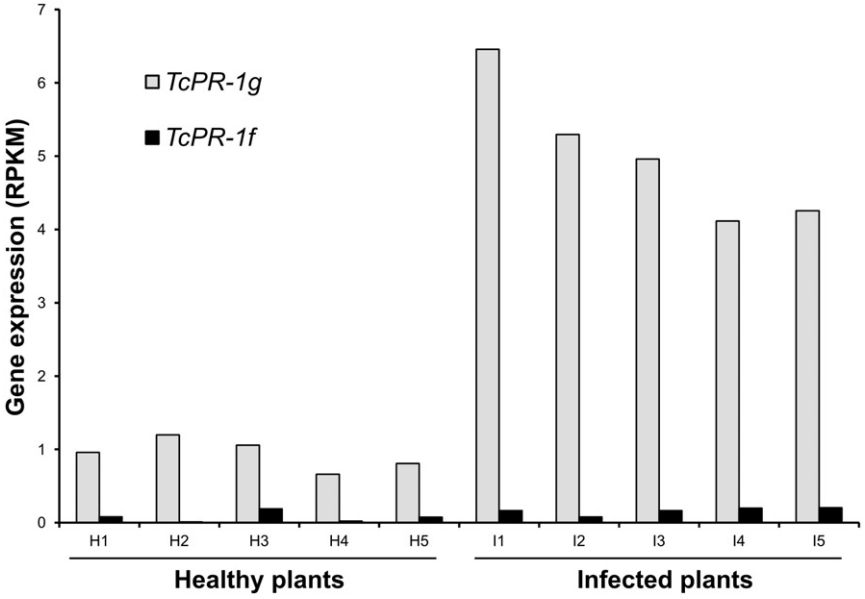


Fig. 3 Gene expression analysis of cacao pathogenesis-related protein 1 receptor-like kinase gene (*PR-1RK*) during the biotrophic stage of witches' broom disease. The *Theobroma cacao TcPR-1g* gene is up-regulated in infected plants, whereas no significant expression was detected for *TcPR-1f*. Data were obtained by RNA-seq sequencing of five replicates for each condition (healthy control and infected plants). Gene expression values are given in reads per kilobase of exon model per million mapped reads (RPKM).

(Shiu *et al*., 2004). These results indicate that PR-1RKs originated from independent events of fusion between PR-1 and kinase domains in different plant species, representing a rare example of convergent evolution of domain architectures (Gough, 2005).

Based on evidence that both PR-1 proteins and RLKs are involved in plant defence responses, we decided to evaluate the transcription of *TcPR-1RK* genes through RNA-seq data derived from the witches' broom disease (WBD) Transcriptome Atlas (P. J. P. L. Teixeira *et al*., unpublished data). WBD, a severe phytopathological problem in the Americas, is caused by the hemibiotrophic basidiomycete *Moniliophthora perniciosa* (Stahel) (Aime and Phillips-Mora, 2005; Meinhardt *et al*., 2008; Purdy and Schmidt, 1996). The infection experiments were performed using cacao seedlings of the variety 'Comum' grown for approximately 3 months in a glasshouse under controlled conditions of temperature (26 °C) and humidity (>80%), and a photoperiod of 12 h. Apical meristems of five plantlets were infected with 30 μL of a *M. perniciosa* basidiospore suspension (10^5 spores/mL), according to Frias *et al*. (1995). Infected tissues were collected 30 days after inoculation, representing the green broom stage of WBD. Green brooms are hyperplastic cacao stems colonized by the biotrophic mycelia of *M. perniciosa* (Meinhardt *et al*., 2008). Five biological replicates were utilized in this experiment and noninfected plants were used as controls. The inspection of the RNA-seq data revealed that *TcPR-1g* is up-regulated in infected plants ($P < 0.001$), whereas *TcPR-1f* is only expressed at low levels in the conditions analysed (Fig. 3). The up-regulation of *TcPR-1g* expression during the green broom stage may indicate a role for this novel receptor in defence responses against pathogens. The *TcPR-1g* expression profile fits the 'receptor swarm hypothesis' (Lehti-Shiu *et al*., 2009), which posits that the expression of RLKs increases after pathogen perception in compatible interactions.

It seems reasonable to speculate that the extracellular domain of PR-1RKs binds to the same ligands as secreted PR-1s. However, the precise function and identity of these putative ligands of plant PR-1 proteins remain elusive. Because PR-1s seem to play fundamental roles in plant immunity, pathogens may have evolved effector proteins to interfere with PR-1 activity. Indeed, it has been shown recently that a candidate effector from the powdery mildew fungus *Blumeria graminis* interacts with a PR-1 protein from barley (Zhang *et al*., 2012). In this regard, we hypothesize that RLKs with a PR-1 extracellular domain may interact with putative ligands of secreted PR-1 proteins, which would trigger a defence response. The identification of PR-1RK ligands and the

downstream signals transduced on the activation of PR-1RKs will illuminate their mechanisms of action and provide clues about the roles of PR-1 proteins in plant immunity.

ACKNOWLEDGEMENTS

This work was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP—project numbers 2009/51018-1, 2009/50119-9 and 2012/07657-2).

REFERENCES

Afoakwa, E.O., Paterson, A., Fowler, M. and Ryan, A. (2008) Flavor formation and character in cocoa and chocolate: a critical review. *Crit. Rev. Food Sci. Nutr.* **48**, 840–857.

Afzal, A.J., Wood, A.J. and Lightfoot, D.A. (2008) Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Mol. Plant–Microbe Interact.* **21**, 507–517.

Aime, M.C. and Phillips-Mora, W. (2005) The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. *Mycologia*, **97**, 1012–1022.

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

Cantacessi, C., Campbell, B.E., Visser, A., Geldhof, P., Nolan, M.J., Nisbet, A.J., Matthews, J.B., Loukas, A., Hofmann, A., Otranto, D., Sternberg, P.W. and Gasser, R.B. (2009) A portrait of the 'SCP/TAPS' proteins of eukaryotes—developing a framework for fundamental research and biotechnological outcomes. *Biotechnol. Adv.* **27**, 376–388.

Chen, X., Shang, J., Chen, D., Lei, C., Zou, Y., Zhai, W., Liu, G., Xu, J., Ling, Z., Cao, G., Ma, B., Wang, Y., Zhao, X., Li, S. and Zhu, L. (2006) A B-lectin receptor kinase gene conferring rice blast resistance. *Plant J.* **46**, 794–804.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J.D., Felix, G. and Boller, T. (2007) A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature*, **448**, 497–500.

Dardick, C., Chen, J., Richter, T., Ouyang, S. and Ronald, P. (2007) The rice kinase database. A phylogenomic database for the rice kinome. *Plant Physiol.* **143**, 579–586.

Diener, A.C. and Ausubel, F.M. (2005) Resistance to *Fusarium oxysporum* 1, a dominant Arabidopsis disease-resistance gene, is not race specific. *Genetics*, **171**, 305–321.

Escobar-Restrepo, J.M., Huck, N., Kessler, S., Gagliardini, V., Gheyselinck, J., Yang, W.C. and Grossniklaus, U. (2007) The FERONIA receptor-like kinase mediates male–female interactions during pollen tube reception. *Science*, **317**, 656–660.

Evans, H.C. (2007) Cacao diseases—the trilogy revisited. *Phytopathology*, **97**, 1640–1643.

Fernández, C., Szyperski, T., Bruyère, T., Ramage, P., Mösinger, E. and Wüthrich, K. (1997) NMR solution structure of the pathogenesis-related protein P14a. *J. Mol. Biol.* **26**, 576–593.

Frias, G.A., Purdy, L.H. and Schmidt, R.A. (1995) An inoculation method for evaluating resistance of cacao to *Crinipellis perniciosa*. *Plant Dis.* **79**, 787–791.

Geer, L.Y., Domrachev, M., Lipman, D.J. and Bryant, S.H. (2002) CDART: protein homology by domain architecture. *Genome Res.* **12**, 1619–1623.

Gibbs, G.M., Roelants, K. and O'Bryan, M.K. (2008) The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins—roles in reproduction, cancer, and immune defense. *Endocr. Rev.* **29**, 865–897.

Gish, L.A. and Clark, S.E. (2011) The RLK/Pelle family of kinases. *Plant J.* **66**, 117–127.

Godiard, L., Sauviac, L., Torii, K.U., Grenon, O., Mangin, B., Grimsley, N.H. and Marco, Y. (2003) ERECTA, an LRR receptor-like kinase protein controlling development pleiotropically affects resistance to bacterial wilt. *Plant J.* **36**, 353–365.

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

Gough, J. (2005) Convergent evolution of domain architectures (is rare). *Bioinformatics*, **21**, 1464–1471.

Greeff, C., Roux, M., Mundy, J. and Petersen, M. (2012) Receptor-like kinase complexes in plant innate immunity. *Front. Plant Sci.* **3**, 209.

Heese, A., Hann, D.R., Gimenez-Ibanez, S., Jones, A.M., He, K., Li, J., Schroeder, J.I., Peck, S.C. and Rathjen, J.P. (2007) The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proc. Natl. Acad. Sci. USA*, **104**, 12217–12222.

Jinn, T.L., Stone, J.M. and Walker, J.C. (2000) HAESA, an Arabidopsis leucine-rich repeat receptor kinase, controls floral organ abscission. *Genes Dev.* **14**, 108–117.

Johnson, L.N., Noble, M.E. and Owen, D.J. (1996) Active and inactive protein kinases: structural basis for regulation. *Cell*, **85**, 149–158.

Jones, C.D., Custer, A.W. and Begun, D.J. (2005) Origin and evolution of a chimeric fusion gene in *Drosophila subobscura*, *D. madeirensis* and *D. guanche*. *Genetics*, **170**, 207–219.

Jurka, J. (1998) Repeats in genomic DNA: mining and meaning. *Curr. Opin. Struct. Biol.* **8**, 333–337.

Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., Minami, E. and Shibuya, N. (2006) Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci. USA*, **103**, 11 086–11 091.

Lee, S.W., Han, S.W., Sririyanum, M., Park, C.J., Seo, Y.S. and Ronald, P.C. (2009) A type I-secreted, sulfated peptide triggers XA21-mediated innate immunity. *Science*, **326**, 850–853.

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

Liu, B., Li, J.F., Ao, Y., Qu, J., Li, Z., Su, J., Zhang, Y., Liu, J., Feng, D., Qi, K., He, Y., Wang, J. and Wang, H.B. (2012) Lysin motif-containing proteins LYP4 and LYP6 play dual roles in peptidoglycan and chitin perception in rice innate immunity. *Plant Cell*, **24**, 3406–3419.

Long, M., Betran, E., Thornton, K. and Wang, W. (2003) The origin of new genes: glimpses from the young and old. *Nat. Rev. Genet.* **4**, 865–875.

Meinhardt, L.W., Rincones, J., Bailey, B.A., Aime, M.C., Griffith, G.W., Zhang, D. and Pereira, G.A. (2008) *Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe? *Mol. Plant Pathol.* **9**, 577–588.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H. and Shibuya, N. (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **104**, 19 613–19 618.

Monaghan, J. and Zipfel, C. (2012) Plant pattern recognition receptor complexes at the plasma membrane. *Curr. Opin. Plant Biol.* **15**, 349–357.

Morillo, S.A. and Tax, F.E. (2006) Functional analysis of receptor-like kinases in monocots and dicots. *Curr. Opin. Plant Biol.* **9**, 460–469.

Nakai, K. and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* **24**, 34–36.

Niderman, T., Genetet, I., Bruyère, T., Gees, R., Stintzi, A., Legrand, M., Fritig, B. and Mösinger, E. (1995) Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans*. *Plant Physiol.* **108**, 17–27.

Nisole, S., Lynch, C., Stoye, J.P. and Yap, M.W. (2004) A Trim5-cyclophilin A fusion protein found in owl monkey kidney cells can restrict HIV-1. *Proc. Natl. Acad. Sci. USA*, **101**, 13 324–13 328.

Oh, M.H., Wang, X., Kota, U., Goshe, M.B., Clouse, S.D. and Huber, S.C. (2009) Tyrosine phosphorylation of the BRI1 receptor kinase emerges as a component of brassinosteroid signaling in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **106**, 658–663.

Oh, M.H., Wang, X., Wu, X., Zhao, Y., Clouse, S.D. and Huber, S.C. (2010) Autophosphorylation of Tyr-610 in the receptor kinase BAK1 plays a role in brassinosteroid signaling and basal defense gene expression. *Proc. Natl. Acad. Sci. USA*, **107**, 17 827–17 832.

Purdy, L.H. and Schmidt, R.A. (1996) Status of cacao witches' broom: biology, epidemiology, and management. *Annu. Rev. Phytopathol.* **34**, 573–594.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. and Lopez, R. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* **33** (Web Server Issue), W116–W120.

Rauscher, M., Adám, A.L., Wirtz, S., Guggenheim, R., Mendgen, K. and Deising, H.B. (1999) PR-1 protein inhibits the differentiation of rust infection hyphae in leaves of acquired resistant broad bean. *Plant J.* **19**, 625–633.

Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., Malinovsky, F.G., Tör, M., de Vries, S. and Zipfel, C. (2011) The Arabidopsis leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required

for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell*, **23**, 2440–2455.

Shiu, S.H. and Bleecker, A.B. (2001) Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proc. Natl. Acad. Sci. USA*, **98**, 10 763–10 768.

Shiu, S.H. and Bleecker, A.B. (2003) Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiol.* **132**, 530–543.

Shiu, S.H., Karlowski, W.M., Pan, R., Tzeng, Y.H., Mayer, K.F. and Li, W.H. (2004) Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. *Plant Cell*, **16**, 1220–1234.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539.

Soria, J.V. (1996) Obtención de clones de cacao por el método de índices de selección. *Turrialba* **16**, 119–124.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739.

Torii, K.U. (2004) Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. *Int. Rev. Cytol.* **234**, 1–46.

Trotochaud, A.E., Jeong, S. and Clark, S.E. (2000) CLAVATA3, a multimeric ligand for the CLAVATA1 receptor-kinase. *Science*, **289**, 613–617.

Van Loon, L.C., Rep, M. and Pieterse, C.M. (2006) Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* **44**, 135–162.

Vello, F. and Garcia, J.R. (1971) Características das principais variedades de cacau cultivadas na Bahia. *Theobroma*, **1**, 3–10.

Walker, J.C. (1994) Structure and function of the receptor-like protein kinases of higher plants. *Plant Mol. Biol.* **26**, 1599–1609.

Wan, J., Zhang, X.C., Neece, D., Ramonell, K.M., Clough, S., Kim, S.Y., Stacey, M.G. and Stacey, G. (2008) A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. *Plant Cell*, **20**, 471–481.

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

Wang, W., Zheng, H., Fan, C., Li, J., Shi, J., Cai, Z., Zhang, G., Liu, D., Zhang, J., Vang, S., Lu, Z., Wong, G.K., Long, M. and Wang, J. (2006) High rate of chimeric gene origination by retroposition in plant genomes. *Plant Cell*, **18**, 1791–1802.

Wang, Z.-Y., Seto, H., Fujioka, S., Yoshida, S. and Chory, J. (2001) BRI1 is a critical component of a plasma-membrane receptor for plant steroids. *Nature*, **410**, 380–383.

Warren, A.M., Hughes, M.A. and Crampton, J.M. (1997) Zebedee: a novel copia-Ty1 family of transposable elements in the genome of the medically important mosquito *Aedes aegypti*. *Mol. Gen. Genet.* **254**, 505–513.

Willmann, R., Lajunen, H.M., Erbs, G., Newman, M.A., Kolb, D., Tsuda, K., Katagiri, F., Fliegmann, J., Bono, J.J., Cullimore, J.V., Jehle, A.K., Götz, F., Kulik, A., Molinaro, A., Lipka, V., Gust, A.A. and Nürnberger, T. (2011) Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. USA*, **108**, 19 824–19 829.

Zhang, W.J., Pedersen, C., Kwaaitaal, M., Gregersen, P.L., Mørch, S.M., Hanisch, S., Kristensen, A., Fuglsang, A.T., Collinge, D.B. and Thordal-Christensen, H. (2012) Interaction of barley powdery mildew effector candidate CSEP0055 with the defence protein PR17c. *Mol. Plant Pathol.* **13**, 1110–1119.

Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E.J., Jones, J.D., Felix, G. and Boller, T. (2004) Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature*, **428**, 764–767.

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