# Involvement of structurally distinct cupuassu chitinases and osmotin in plant resistance to the fungus *Moniliophthora perniciosa*

Raner José Santana Silva[1], Rafael Moyses Alves[2], Karina Peres Gramacho[3], Lucilia Helena Marcellino[4], Fabienne Micheli[1,5,*]

[1]Universidade Estadual de Santa Cruz (UESC), Departamento de Ciências Biológicas (DCB), Centro de Biotecnologia e Genética (CBG), Rodovia Ilhéus-Itabuna, km 16, 45662-900 Ilhéus-BA, Brazil.

[2]Empresa Brasileira de Pesquisa Agropecuária, Embrapa Amazônia Oriental, 66095-903 Belém-PA, Brazil.

[3]Cocoa Research Center, CEPLAC/CEPEC, 45600-970 Itabuna-BA, Brazil.

[4]Empresa Brasileira de Pesquisa Agropecuária, Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, 70770-917, Brazil.

[5]CIRAD, UMR AGAP, F-34398 Montpellier, France.

***Corresponding author:** Dr Fabienne Micheli, UESC, DCB, Rodovia Ilhéus-Itabuna km16, 45662-900, Ilhéus-BA, Brazil. Phone: +55 73 3680 5196. Fax: +55 73 3680 5226. E-mail: fabienne.micheli@cirad.fr

**Abstract:**

The cupuassu tree (*Theobroma grandiflorum*) is a crop of great economic importance to Brazil, mainly for its pulp and seeds, which are used in food industry. However, cupuassu fruit production is threatened by witches' broom disease caused by the fungus *Moniliophthora perniciosa*. As elements of its defense mechanisms, the plant can produce and accumulate

pathogenesis-related (PR) proteins such as chitinases and osmotins. Here, we identified three PR proteins from cupuassu (TgPR3, TgPR5 and TgPR8) from cupuassu-*M. perniciosa* RNA-seq data. TgPR3 and TgPR8 corresponded to chitinases, and TgPR5 to osmotin; they are phylogenetically related to cacao and to Arabidopsis PR sequences involved in biotic and abiotic stress. The TgPR proteins' tridimensional structure was obtained through homology modeling, and molecular docking with chitin and chitosan showed that the TgPR proteins can interact with both cell wall molecules and presented a higher affinity for chitosan. TgPR gene expression was analyzed by RT-qPCR on resistant and susceptible cupuassu genotypes infected by *M. perniciosa* at 8, 24, 48 and 72 hours after infection (hai). The TgPR genes showed higher expression in resistant plants compared to the susceptible ones, mainly for TgPR5 at 8 and 24 hai, while the expression was lower in the susceptible cupuassu plants. To our knowledge, this is the first *in silico* and *in vitro* reports of cupuassu PR protein. The data suggested that TgPRs could be involved in recognizing mechanisms of the plant's innate immune system through chitin receptors. Our results also suggest a putative role of chitinase/chitosanase for the TgPR5/osmotin.

**Keywords:** Pathogenesis-related proteins, molecular docking, witches' broom disease, gene expression, chitin, chitosan

## 1. Introduction

The cupuassu (*Theobroma grandiflorum* [Willd. Ex Spreng.] Schum.) belongs to the Malvaceae family; it is native to the Amazon region and is economically important to Brazil, where it is cultivated [1, 2]. Cupuassu pulp and seeds are used in the cosmetics and food industries, mainly for candy, ice cream, beverage and *cupulate* (a product similar to chocolate) production [3]. *Cupulate* is an interesting alternative to chocolate production for an

actual worldwide economic situation in which the production of cocoa (*Theobroma cacao*, Malvaceae family) beans cannot supply the increased worldwide demand for chocolate [4, 5]. However, both cocoa and cupuassu cultures are affected by pests and pathogens that cause damage, loss of production and even plant death. Among them, witches' broom disease, caused by the fungus *Moniliophthora perniciosa*, is the main phytopathological problem for cupuassu cultures [6].

Because they are attacked by a wide variety of pathogens, plants have developed pattern-recognition receptors to recognize molecular patterns associated with pathogens (PAMPs), which lead, among other outcome, to the production and accumulation of pathogenesis-related proteins (PR proteins), which are considered crucial to plant defense mechanisms [7-9]. The current classification divides PR proteins into 17 classes (PR1–PR17) depending on their biological activity, physicochemical properties and/or sequence homology [10, 11]. The PR-3, PR-4, PR-8 and PR-11 families correspond to different types of chitinases (EC 3.2.1.14). Among them, the PR-3 proteins include glycoside hydrolase 19 (GH19) chitinases, and the PR-8 family corresponds to glycoside hydrolase 18 (GH18) chitinases. The PR-5 family classification corresponds to thaumatin-like protein (TLPs), whose biological function in plants has yet to be established [12]. Among TLPs, the osmotins are multifunctional proteins that may participate alongside chitinases in protein complexes involved in plant defense against pathogens [8, 13]. GH18, GH19 and osmotin show antifungal activity through cell wall hydrolysis or plasma membrane permeabilization [12, 14]; in the case of PR-5, it has been proposed that the existence of specific target receptors on the membrane leads to fungus sensitivity or resistance [15]. The GH18 chitinases show a characteristic barrel structure [16], while GH19 chitinases are composed of a high concentration of α-helices and loops [14]. The TLP proteins have three well-characterized domains: domain I is composed of β sheets and loops called a β-sandwich, domain II is

composed of α-helices and domain III is composed of two β-sheets [17]. Because chitin is present in the fungus cell wall and not in plant cells, it is considered an ideal PAMP and as one of the first barriers in the plant-pathogen battlefield [18, 19]. Chitin and chitosan, its deacetylated derivative, have been heavily studied, and their role in plant–pathogen interaction has received considerable attention, mainly because they are considered useful for biotechnological applications aimed at reducing fungus activity and/or development [19-21]. *Moniliophthora perniciosa*'s chitin metabolism, including associated enzymes such as chitin synthases and chitinases, has previously been studied [22-25].

Molecular studies of cupuassu related to plant–pathogen interaction are still poorly developed, and few works are related to resistance to witches' broom disease [26-28]. Moreover, unlike cocoa, the cupuassu genome has not been sequenced yet, and molecular analyses focused on genetics and genomics of resistance are only based on recent transcriptomic data [26, 28]. Here, we selected and analyzed three PR genes from cupuassu (TgPR3, TgPR5 and TgPR8) with putative chitinase or chitinase-associated functions as well as anti-fungal activity. We showed, *in silico*, the TgPR proteins' structural aspects and potential activity in relation to chitin and chitosan ligands, and, *in vitro*, the contrasting expression levels of the *TgPR* genes in resistant and susceptible cupuassu plants inoculated with *M. perniciosa*. Their expression profiles could be related to phylogenetic analysis using homologous PR sequences from *T. cacao* and *A. thaliana*, which are involved in response to biotic and abiotic stress. To our knowledge, this is the first PR protein report on cupuassu–*M. perniciosa* interaction. The overall data data suggest that TgPRs could be involved in recognizing mechanisms of the plant's innate immune system through chitin receptors. Our results also suggest a putative role of chitinase/chitosanase for the TgPR5/osmotin.

## 2. Methods

101 **2.1 Identification and characterization of *TgPR* sequences**

102 The PR sequences were selected from a previously obtained RNA-Seq database of cupuassu
103 [28]. Open reading frame (ORF) detection was performed using the ORFinder software
104 (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The proteins were analyzed with BLASTp
105 against the Conserved Domain Database (CDD;
106 https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) for conserved domain identification,
107 protein sequence length checking and TgPR protein classification. Signal peptide was
108 identified using the SignalP 4.1 software [29], while the molecular weight and isoelectric
109 point were obtained using the Compute pI/Mw program [30]. The subcellular locations were
110 predicted using the YLoc-HighRes Plants tool [31].

111

112 **2.2 Phylogeny**

113 Cacao and Arabidopsis nucleotide sequences that are homologous to TgPRs were obtained
114 using local alignment (BLAST) from CocoaGenDB [32] and TAIR
115 (https://www.arabidopsis.org), respectively. The global multiple alignment of all sequences
116 was made using ClustalW available with MEGAX software [33]. The MEGAX software was
117 used to construct a rooted phylogenetic tree using the maximum likelihood method with 1,000
118 bootstrap sampling steps, the Tamura-Nei model and the G+I substitution method.
119 Phylogenetic trees were visualized and edited using the FigTree v1.4.4 tool
120 (http://tree.bio.ed.ac.uk/software/figtree/).

121

122 **2.3 Homology modeling of TgPRs and molecular docking with chitin and chitosan**

123 The tri-dimensional (3D) structure of the TgPR proteins was obtained using the Swiss-Model
124 server [34]. The crystal structures of homologous proteins available from the Protein Data
125 Bank (PDB; https://www.rcsb.org/) were used as templates to build the structural models of

126 TgPR (Supplementary material 1). The stereochemical quality of the models was analyzed
127 with ANOLEA [35] and Procheck [36]. Water molecules were removed from and polar
128 hydrogens were added to the TgPR protein models, before calculation of their Gasteiger
129 charges; the results were exported in PDBQT format. The active sites were identified based
130 on TgPR protein alignment with PR proteins from other species (Supplementary materials 2
131 to 4), using the ClustalW tool [37]. The amino acid residues from the TgPR active sites of the
132 proteins as well as the grid box defining the docking region were marked using AutoDock
133 Tools v1.5.6 [38]. For docking, chitin (CID: 444514) and chitosan (CID: 71853) were used as
134 ligands; their structures were downloaded in SMILES format. Using the MarvinSketch
135 15.7.13.0 tool, the ligands were converted into 3D structures, checked for conformation and
136 saved in mol2 format. The Kollman charges were calculated using AutoDockTools v1.5.6,
137 and the results were exported in PDBQT format. The docking was done using the AutoDock
138 Vina software [39] with default parameters. The results were visualized in PyMOL v1.7.4
139 [40] and Discovery Studio 4.5.
140

141 **2.4 Plant material**

142 *TgPR* gene expression was analyzed in adult plants obtained from clonal cuttings of cupuassu
143 genotypes C174 (resistant to witches' broom disease) and C1074 (susceptible to witches'
144 broom disease), grown in the experimental station of CEPLAC (Belém, Pará, Brazil) [41].
145 Cupuassu apical meristems were inoculated with *M. perniciosa* using the droplet method (30
146 µL of suspension containing $1x10^5$ $ml^{-1}$ basidiospores) [42]. After inoculation, the stem apex
147 region was wrapped in a plastic bag (a water-saturated environment) to improve basidiospore
148 germination and plant infection by the fungus [43]. The control plants were submitted to the
149 same procedure, but the basidiospore suspension was replaced by distilled water. The apical
150 meristems of the inoculated C174 and C1074 genotypes were harvested 8, 24, 48 and 72

151 hours after inoculation (hai), while non-inoculated ones (control) were harvested at 8 hai
152 Three plants (biological replicates) were used for each genotype at each harvesting time. The
153 harvested samples were immediately frozen in liquid nitrogen and stored at −80 °C until use.
154

155 **2.5 Reverse transcription quantitative PCR (qPCR) analysis**

156 Each biological replicate (individual plant) was macerated in liquid nitrogen, and 3 mg of the
157 macerate was used to extract its total RNA using the RNAqueous® kit according to the
158 manufacturer's recommendations (Ambion). Total RNA (15 µl) was treated with the DNase I
159 kit according to the manufacturer's recommendations (Invitrogen). The cDNA was
160 synthesized using the RevertAid First Strand cDNA Synthesis Kit, a cDNA synthesis kit,
161 according to the manufacturer's recommendations (ThermoScientific). The cDNA was
162 quantified on a GeneQuant Pro Spectrophotometer UV/Vis Reader (Amersham). For qPCR,
163 the *TgPR* primers were designed with the Primer 3 plus tool [44] using previously described
164 parameters [45], and checked for dimer and hairpin formation using the Oligoanalyzer
165 (https://www.idtdna.com/calc/analyzer) (Supplementary material 5). Three reference genes
166 (acyl carrier protein/ACP, malate dehydrogenase/MDH and tubulin/TUB; Supplementary
167 material 5) previously identified in cupuassu [27] were analyzed for stability in three pooled
168 samples of each cupuassu genotype: i) a non-inoculated pool sample (3 plants); ii) a pool of 8
169 and 24 hai samples (6 plants); and iii) a pool of 48 and 72 hai samples (6 plants). Quantitative
170 PCR was performed in an Eppendorf thermocycler Realplex4. The reaction was made in a
171 final volume of 10 µl containing 5 µl of Taq READYMIX SYBR Green, 0.1 µl of ROX
172 (fluorescence signal normalizer), 0.5 µl of primers (forward and reverse, at a concentration of
173 10 mM each) and 4.4 µl of cDNA 1/20 (230 ng). The cycling conditions were 95°C for 2 min
174 followed by 45 cycles of 95°C for 15 s, 55°C for 15 s and 60°C for 45 s, followed by a
175 dissociation curve step. The fluorescence data and primer efficiency were analyzed using

176 MINER 4.0 [46]. The gene-expression stability was calculated using NormFinder [47], which
177 determined the best combination of the reference genes ACP and MDH for relative expression
178 calculation (Supplementary material 6). The gene expression was analyzed in each biological
179 replicate (1 biological replicate = 1 plant) for each genotype (inoculated vs. non-inoculated) at
180 each harvesting time, using the qPCR cycling conditions described above and three technical
181 replicates. The relative expression was obtained using the REST-2009 software [48], which
182 considers the primer efficiency as well as the average ACP and MDH values for calculation.
183

184 **3. Results**

185 **3.1 Identification and characterization of *TgPR* sequences**

186 Analysis of the cupuassu RNA-seq allowed the selection of three sequences – codified as
187 C106, C68 and C356 – and showed homology with chitinases [28]. Sequence C106 showed
188 homology with chitinases from the GH19 family, which are involved in pathogenesis and are
189 related to the PR3 class (Table 1). For this reason, sequence C106 was re-named as *T.*
190 *grandiflorum* pathogenesis-related protein 3 (TgPR3). The *TgPR3* ORF was 966 bp in length
191 and encoded for a 322 aa protein with a molecular weight and an isoelectric point of 35.5 kDa
192 and 6.64, respectively (Table 1). TgPR3 contained a signal peptide (26 aa) responsible for
193 extracellular addressing (86.48% of probability). The protein contained 17 and three predicted
194 phosphorylation and glycosylation sites, respectively (Table 1). The C68 sequence showed
195 homology with the thaumatin like protein (TLP-P) family, which belongs to the pathogenesis-
196 related protein 5 (PR5) class and includes chitinases and osmotins. The C68 sequence was
197 renamed *T. grandiflorum* pathogenesis-related protein 5 (TgPR5) (Table 1). The TgPR5's
198 ORF was 672 bp in length, and it encoded a 224 aa protein with a molecular weight of 24.2
199 kDa and an isoelectric point of 6.68 (Table 1). The protein contained a 23 aa signal peptide
200 responsible for extracellular addressing (99.7% probability). TgPR5 contains one predicted

201 glycosylation and eight predicted phosphorylation sites (Table 1). The protein encoded by the
202 C356 sequence was classified as a chitinase from the GH18 family and as hevamine enzyme,
203 which are chitinases belonging to the class of pathogenesis-related protein 8 (PR8). This
204 sequence was renamed *T. grandiflorum* pathogenesis-related protein 8 (TgPR8). The TgPR8
205 ORF was 927 bp in length, and it encoded for a 309 aa protein with a molecular weight of
206 33.3 kDa and a pI of 4.27 (Table 1). TgPR8 contains a signal peptide (29 aa), which
207 potentially addresses the protein to the extracellular (52.2% probability) or the vacuolar
208 (34.5% probability) compartments. The TgPR8 protein showed one and 13 glycosylation and
209 phosphorylation sites, respectively (Table 1).
210
211 **3.2 Phylogeny**
212 BLAST made from *TgPR* sequences allowed the identification of 11 and 14 homologous
213 sequences of TgPR3, 30 and 28 of TgPR5 and 14 and one of TgPR8 in *T. cacao* and *A.*
214 *thaliana*, respectively (Fig. 1). TgPR3 showed a high homology with *T. cacao* sequences
215 Tc04_t029180 and Tc06_t000490 and with *A. thaliana* sequences AT3G16920.1 and
216 AT1G05850.1 (Fig. 1A). TgPR5 was clustered with the sequences Tc00_t060970,
217 Tc03_t026960, Tc03_t026980, Tc03_t026990, Tc03_t027000, Tc03_t027010 and
218 Tc03_t027030 from *T. cacao* and with AT4G11650 (OSM34) from *A. thaliana* (Fig. 1B).
219 TgPR8 was clustered with the sequences Tc01_t032120, Tc03_t017760, Tc03_t017780,
220 Tc03_t017790, Tc10_t015260 and Tc10_t015330 from *T. cacao* as well as with AT5G24090
221 from *A. thaliana* (Fig. 1C).
222
223 **3.3 Homology modeling and validation of TgPR's 3D structure**
224 The TgPR3 protein's 3D structure (Fig. 2A) was obtained through homology modeling with
225 the GH19 chitinase (PDB: 4TX7) from *Vigna ungiculata* (39% of identity and 91% of

226 coverage; Supplementary material 1). The TgPR3 protein 3D model showed that 87% of the
227 amino acid residues was in the most favored regions, 11.8% was in the additional allowed
228 regions, 0.6% was in the generously allowed regions and 0.6% was in the disallowed regions
229 (Supplementary material 7A). The TgPR3 model contained 12 helices, 26 β-turns, eight ɣ-
230 turns and three disulfide bonds (Fig. 2A, Supplementary material 8). The alignment of TgPR3
231 with other GH19 chitinases from other species (Supplementary material 2) allowed the
232 identification of TgPR3's catalytic site containing the amino acid residues $Lys_{70}$, $Glu_{92}$ and
233 $Tyr_{125}$ (Fig. 2A). The TgPR5 protein's 3D structure (Fig. 2B) was obtained by homology
234 modeling with an osmotin (PDB: 4L2J) from *Calotropis porcine* (75% of identity and 100%
235 of coverage; Supplementary material 1). The TgPR5 protein 3D model showed 88.5% of the
236 amino acid residues in the most favored regions, 11.1% in the generously allowed regions and
237 0.4% in the allowed regions, and it had no amino acid residues in disallowed regions
238 (Supplementary material 7B). The TgPR5 model contained 14 β-sheets, five helices, 25 β-
239 turns, two ɣ-turns, one bulge, five hairpins and seven disulfide bonds (Fig. 2B, Supplementary
240 material 8). The alignment of TgPR5 protein with thaumatine-like proteins from other species
241 (Supplementary material 3) allowed the identification of putative amino acid residues from
242 the active cleft: $Arg_{43}$, $Glu_{83}$, $Asp_{96}$ and $Asp_{101}$ (Fig. 2B). The TgPR8 protein's 3D structure
243 (Fig. 2C) was obtained by homology modeling with hevamine A (PDB: 1LLO) from *Hevea*
244 *brasiliensis* (64% of identity and 99% of coverage; Supplementary material 1). The TgPR8
245 protein's 3D structure showed 78.8% of the amino acid residues in the most favored regions,
246 18.9% in the additional allowed regions, 0.9% in the generously allowed regions and 1.4% in
247 disallowed regions (Supplementary material 7C). The TgPR8 model contained a barrel shape
248 with nine β-sheets, 15 helices, three disulfide bonds, 16 β-turns, one ɣ-turn, four bulges, one
249 hairpin, six α-β units and one PSI loop (Fig. 2C, Supplementary material 8). The alignment of
250 TgPR8 with GH18 chitinases from other species (Supplementary material 4) allowed the

251 identification of motifs and conserved domains, including the predicted active site containing
252 the amino acid residues $Tyr_{6}$, $Phe_{32}$, $Asp_{125}$, $Glu_{127}$, $Gln_{155}$, $Gln_{178}$, $Tyr_{180}$ and $Trp_{253}$ (Fig. 2C).
253

254 **3.4 Molecular docking between TgPRs and chitin or chitosan**

255 The molecular docking of TgPR3 with chitin had an affinity energy of −6.4 kcal/mol
256 (Supplementary material 9). It showed conventional hydrogen bond interactions with amino
257 acid residues $Ser_{92}$, $Gly_{94}$, $Glu_{112}$ and $Tyr_{225}$, and carbon hydrogen bonds with amino acid
258 residues $Lys_{90}$ and $Pro_{217}$ (Fig. 3). The molecular docking of TgPR3 with chitosan had an
259 affinity energy of −7.4 kcal/mol (Supplementary material 9). It showed conventional
260 hydrogen bond interactions with amino acid residues $Lys_{90}$, $Ser_{92}$, $Asp_{118}$, $Tyr_{119}$, $Tyr_{145}$,
261 $Tyr_{148}$, $Arg_{182}$, $Leu_{224}$, $Tyr_{225}$ and $Asp_{236}$; carbon hydrogen bonds with amino acid residues
262 $Thr_{91}$, $Tyr_{109}$ and $Arg_{111}$; and alkyl interactions with the amino acid residue $Pro_{143}$ (Fig. 3).
263 The molecular docking of TgPR5 with chitin had an affinity energy of −6.5 kcal/mol
264 (Supplementary material 9) and showed conventional hydrogen bonds with the amino acid
265 residues $Thr_{40}$, $Glu_{83}$, $Gln_{88}$, $Asp_{96}$ and $Pro_{179}$; one carbon hydrogen bond with the amino acid
266 residue $Asp_{182}$; and one pi-sigma interaction with the amino acid residue $Tyr_{176}$ (Fig. 3). The
267 molecular docking of TgPR5 with chitosan had an affinity energy of −6.8 kcal/mol
268 (Supplementary material 9) with conventional hydrogen bonds in the amino acid residues
269 $Thr_{40}$, $Ala_{42}$, $Tyr_{74}$, $Gly_{75}$, $Glu_{83}$, $Tyr_{84}$, $Asp_{96}$, $Asp_{101}$, $Tyr_{176}$ and $Asp_{182}$ and one carbon
270 hydrogen bond in the amino acid residue $Cys_{154}$ (Fig. 3). The molecular docking of TgPR8
271 with chitin had an affinity energy of −7.0 kcal/mol (Supplementary material 9) and showed
272 conventional hydrogen bonds with the amino acid residues $Gln_{9}$, $Ala_{81}$, $Glu_{127}$, $Gln_{155}$, $Tyr_{180}$
273 and $Tyr_{257}$ (Fig. 3). The docking of TgPR8 with chitosan had an affinity energy of −7.3
274 kcal/mol (Supplementary material 9) and showed conventional hydrogen bonds with the

275 amino acid residues $Gln_9$, $Ser_{43}$, $Met_{44}$, $Asn_{45}$, $Leu_{46}$, $Gln_{155}$, $Asn_{181}$, $Asn_{182}$, $Ala_{221}$, $Trp_{253}$ and
276 $Tyr_{257}$, and carbon hydrogen bonds with the amino acid residues $Gln_{178}$ and $Tyr_{180}$ (Fig. 3).
277
278 **3.5 Analysis of *TgPR* gene expression**
279 The primers PR3, PR5, PR8, ACP and MDH showed efficiency from 91% (MDH on the
280 C1074 genotype) to 102% (P3 on both the C174 and C1074 genotypes; Supplementary
281 material 10). For both genotypes and for all of the harvesting points, the PCR amplification
282 occurred at the same and unique melting temperature for each gene and a unique band was
283 visible on electrophoresis agarose gel, showing that only the corresponding gene was
284 amplified (Supplementary material 11). The relative expression of the *TgPR* genes was lower
285 in the susceptible genotype at all of the harvesting times, as compared to the resistant
286 genotype (Figs. 4A and B). The *TgPR3* gene showed two to three times higher expression in
287 the resistant genotype than in the susceptible one (Fig. 4B). Similarly, the *TgPR8* gene
288 showed two to four times higher expression in the resistant plants compared to the susceptible
289 ones (Fig. 4B). For both the *TgPR3* and *TgPR8* genes, the expression through the time course
290 disease slightly increased from 8 to 72 hai in the resistant plants (Fig. 4B). The *TgPR5* gene
291 showed a very high expression at 8 and 24 hai in the resistant genotype (about 25 of relative
292 expression; Fig. 4A), which decreased at 48 and 72 hai (about 2 of relative expression; Fig.
293 4B). Except for *TgPR5* 8 hai, the relative expression of the *TgPR* genes in the susceptible
294 genotype was almost constant (Fig. 4B).
295
296 **4. Discussion**
297 **4.1 TgPRs are structurally distinct and show a high energy affinity for chitosan**
298 The TgPR3 and TgPR5 proteins showed similar isoelectric points and share the same cellular
299 addressing (extracellular; Table 1), whereas the TgPR8 protein shows a more acidic

300 isoelectric point. Although the predictions for TgPR8 suggested extracellular and vacuolar
301 addressing (Table 1), type GH18 chitinases are found in the apoplast [49]. Although GH19
302 (PR-3) chitinases are composed of α-helices and loops, they are highly thermostable [50].
303 GH19 are divided into loopful and loopless classes, where loopful chitinases have loops at
304 both ends of the substrate-binding site, while loopless chitinases have a non–end loop
305 structure, which is commonly found in bacteria [51]. The TgPR3 protein showed homology
306 with the loopful GH19 class because of the presence of the loops near the substrate-binding
307 site (Fig. 2A). Most GH18 proteins (PR-8) correspond to chitinases, although some have been
308 reported to have xylanase-inhibiting activity or to have lost their chitinase activity [16]. As
309 chitinases, the GH18 proteins have the same activity as GH19 but are extremely distinct, both
310 structurally and evolutionarily [52]. One pattern that defines the chitinase activity in GH18 is
311 the presence of a DXDXE motif [53]; the TgPR8 showed the DFDIE motif at positions 123–
312 127 of the sequence, corroborating its classification and potential chitinase activity
313 (Supplementary material 8). Osmotin antifungal activity is associated with the presence of the
314 acidic cleft region [15, 54], which in TgPR5 protein is composed of the residues $Glu_{83}$, $Asp_{96}$
315 and $Asp_{101}$ (Figure 1B). *In vitro* studies have shown that barley seeds' GH19 chitinase can
316 efficiently catalyze tetrameric N-acetylglucosamine in dimers but that most GH19 chitinases
317 prefer larger substrates [55]. These data corroborate our docking results, where the TgPR3
318 protein had an affinity of −6.4 kcal/mol with the chitin molecule (the smaller one) and −7.4
319 kcal/mol with the chitosan molecule (the larger one), showing that TgPR3 also has a higher
320 affinity for larger substrates (Supplementary material 9, Fig. 3). TgPR8 protein has similar
321 affinity energy for chitin and chitosan ligands (−7.0 kcal/mol and −7.3 kcal/mol, respectively;
322 Supplementary material 9). The docking results reinforced the putative TgPR8 protein
323 chitinase activity, with an interaction within the DFDIE motif, in which it is essential for the
324 chitinase activity of this protein class [53]. Although structurally different and having

325 different catalytic mechanisms [52], the proteins TgPR3 and TgPR8 have a very close affinity
326 energy for the chitosan molecule. This may be related to their activity type, as both are
327 glycoside hydrolases [56]. The TgPR5 protein had an affinity energy of −6.5 kcal/mol with
328 the chitin molecule and −6.8 kcal/mol with the chitosan molecule (Supplementary material 9;
329 Fig. 3). These results show that TgPR5 can bind chitin molecules and may have potential
330 chitinase and antifungal activity.
331

332 **4.2 TgPRs are homologous to other plant PR proteins involved in biotic and abiotic**
333 **stress**

334 TgPR3 was clustered with Tc06_t000490 and Tc04_t029180 (Fig. 1), which showed a high
335 constitutive expression across treatments, Tc06_t000490 being up-regulated during cacao-
336 *Phytophthora palmivora* interaction [57]. The AT3G16920 from *A. thaliana* that also
337 clustered with TgPR3 (Fig. 1) is involved in response to mechanical stress [58] as well as to
338 heat, salt and drought stress [59]. AT1G05850 is related to cellulose biosynthesis [60] and
339 also to response to high temperatures, salt and drought stress [59]. Sequences Tc00_t060970,
340 Tc03_t026990, Tc03_t027000, Tc03_t027010 and Tc03_t027030, which were clustered with
341 TgPR5 (Fig. 1), showed increased expression levels during cacao interaction with the
342 pathogenic fungi *P. palmivora* and *Colletotrichum theobromicola* [57]. The *A. thaliana*
343 AT4G11650 sequence showed an increased expression with *Agrobacterium tumefaciens* at 24
344 hai [61]. This gene is also regulated by the ethylene and jasmonic acid–signalization pathway
345 [62]. Transgenic Arabidopsis plants knocked-out for the transcription factor WRKY33 –
346 which is mediated by ethylene and the jasmonic acid signalization pathway – showed a
347 susceptibility pattern to *Botritys cinerea* and a reduction of AT4G11650 expression level,
348 while the wild-type plants showed increased AT4G11650 expression in response to the
349 pathogen [62]. The *A. thaliana* sequence AT5G24090, homologous to TgPR8 (Fig. 1),

350 showed an expression pattern specific to stress response; however, no expression was
351 observed during the plant development stage [63]. In *T. cacao*, the Tc03_g017760 gene
352 homologous with TgPR8 (Fig. 1) showed an increased expression level when the plant was
353 inoculated by *P. palmivora* or *Colletotrichum theobromicola* [57]. The PR gene's expression
354 in response to biotic and abiotic stress could be related to the phytohormone pathways that
355 regulate their expression. This could be related to the literature data about *M. perniciosa*,
356 which modifies the metabolism of *T. cacao* during compatible interactions through the
357 phytohormone pathway [64].
358
359 **4.3 *TgPR* genes were highly expressed in the cupuassu genotype resistant to *M.***
360 ***perniciosa***
361 The relative expression of *TgPR3*, *TgPR5* and *TgPR8* was highly significant in the resistant
362 genotype infected by *M. perniciosa*, at all of the harvesting points, but mainly at 8 and 24 hai
363 for *TgPR5* (Fig. 3A). In the *T. cacao–M. perniciosa* interaction, which is closely related to
364 and also deeper studied than the cupuassu-*M. perniciosa* pathosystem, the basidiospore
365 germination and the beginning of the germination tube in the plant tissues were observed at 6
366 hai [65], a time period that is coherent with the high and early expression of the TgPRs (Figs.
367 3A and B). The very high expression observed for TgPR5 since 8 hai would corroborate other
368 reports, in which the PR5 was overexpressed in resistant plant genotypes after pathogen
369 infection [17, 66-68]. The osmotin's function and mode of action are still unclear, and
370 different functions have been shown, such as in inhibiting beta-glucanase or xylanase [17,
371 69]. Here, we propose that TgPR5 may act as chitinase due to its strong *in silico* interaction
372 with chitin and chitosan, as well as its extracellular localization (Fig. 2, Table 1). Thus, we the
373 resistant cupuassu genotype may have been able to recognize *M. perniciosa* throught its
374 innate immune system – probably through effectors [70] – and to generate a defense response

375 against the pathogen from the beginning of the infection process (Fig. 5), which explains why
376 the C174 resistant genotype does not present symptoms or disease features in the field [71]. In
377 the susceptible cupuassu genotype, the TgPR genes showed reduced expression at all of the
378 harvesting times (Fig. 3) showing that the pathogen can bypass the innate immune system of
379 cupuassu and cause virulence. In *T. cacao–M. perniciosa* interaction, the fungus can interfere
380 with the host's metabolism and alter its gene expression [64]. The reduction of plant chitinase
381 expression levels during the cupuassu–*M. perniciosa* interaction can be highly beneficial to
382 the pathogen, which can grow and colonize the plant tissues (Fig. 5). Moreover, the reduction
383 of cell wall degradation by chitinases (i.e. cell wall fragments) reduces the recognition by the
384 plant's innate immune system through chitin receptors [72, 73] (Fig. 5). In contrast, increased
385 plant chitinase expression and secretion can cause serious damage to the pathogen, allowing
386 the plant to recognize the pathogen, and triggering defense mechanisms and the plant's
387 resistance [73] (Fig. 5).
388

389 **5. Conclusion**

390 Here, we have shown the first *in silico* and *in vitro* analysis of pathogenesis-related proteins
391 from *T. grandiflorum*. We show that TgPR3 and TgPR8 are putative chitinases with distinct
392 structures, and that TgPR5 is a putative osmotin, all with affinity for chitin and chitosan
393 molecules. These results suggest a new role for TgPR5/osmotin as chitinase/chitosanase in
394 plant–pathogen interactions. TgPR expression highly increased in the resistant cupuassu
395 genotype infected *by M. perniciosa* (mainly TgPR5 at the first infection times) and was low in
396 the susceptible one. These results could be related to the phylogenetic analysis of the *TgPR*
397 genes showing that *T. grandiflorum* sequences clustered with cacao and Arabidopsis *PR*
398 sequences are involved in response to biotic and abiotic stress, some through phytohormone-

399 signaling pathways. These observations suggest that TgPRs could be involved in recognizing
400 mechanisms of a plant's innate immune system through chitin receptors.
401
402 **Availability of supporting data**
403 The data sets supporting the results of this article are included within the article and its
404 additional files.
405
406 **Abbreviations:** 3D: tridimensional; PAMP: molecular patterns associated with pathogens;
407 PR: pathogenesis related; TLP: thaumatin-like protein
408
409 **Competing interest**
410 No conflicts of interest to declare.
411
412 **Authors' contribution**
413 RJSS was responsible for the execution of all the experimental steps. RJSS, LHM and FM
414 analyzed and discussed the data. RJSS and FM wrote the manuscript. KPG and RMA were
415 responsible for plant material production and inoculation with *M. perniciosa*. RMA, LHM,
416 FM and KPG were responsible for the conception and design of the experiments. FM was
417 responsible for the financial support of the research. FM and LHM were responsible for the
418 advising of RJSS.
419
420 **Acknowledgements**
421 The work of RJSS was supported by the Fundação de Amparo à Pesquisa do Estado da Bahia
422 (FAPESB), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and
423 Embrapa Recursos Genéticos e Biotecnologia (Cenargen). This work was supported by

FAPESB project DTE0038/2013. KPG and FM received a Productivity grant from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The authors thank Dr. Akyla Maria Martins Alves for advices during manuscript writing. This work was made in the frame of the International Consortium in Advanced Biology (CIBA).

## References

[1] R.M. Alves, M.D.V.d. Resende, B.d.S. Bandeira, T.M. Pinheiro, D.C.R. Farias, Avaliação e seleção de progênies de cupuaçuzeiro (*Theobroma grandiflorum*), em Belém, Pará, Revista Brasileira de Fruticultura, 32 (2010) 204-212.

[2] R.M. Alves, C.R.S. Silva, M.S.C. Silva, D.C.S. Silva, A.M. Sebbenn, Diversidade genética em coleções amazônicas de germoplasma de cupuaçuzeiro [*Theobroma grandiflorum* (Willd. ex Spreng.) Schum.], Revista Brasileira de Fruticultura, 35 (2013) 818-828.

[3] M.C. Costa, G.A. Maia, M.d.S.M. Souza Filho, R.W.d. Figueiredo, R.T. Nassu, J.C.S. Monteiro, Conservação de polpa de cupuaçu [*Theobroma grandiflorum* (Willd. Ex Spreng.) Schum] por métodos combinados, Revista Brasileira de Fruticultura, 25 (2003) 213-215.

[4] R. Sayid, Chocolate Could Run Out In 2020 Due To Worldwide Shortage of Cocoa, in: The Daily Mirror online, 2013, pp. http://www.mirror.co.uk/news/world-news/chocolate-could-run-out-2020-2913505.

[5] A. Wexler, World's sweet tooth heats up cocoa. Growing demand from emerging markets is pushing up prices for key ingredient in chocolate, in: The Wall Street Journal, 2014, pp. http://online.wsj.com/news/articles/SB10001424052702304703804579381234107085804.

[6] R.M. Alves, M.D.V. Resende, Avaliação genética de individuos e progênies de cupuaçuzeiro no estado do Para e estimativas de parâmetros genéticos, Revista Brasileira de Fruticultura, 30 (2008) 696-701.

[7] J. Sels, J. Mathys, B.M.A. De Coninck, B.P.A. Cammue, M.F.C. De Bolle, Plant pathogenesis-related (PR) proteins: a focus on PR peptides, Plant Physiology and Biochemistry, 46 (2008) 941-950.

[8] L.C. Van Loon, Induced resistance in plants and the role of pathogenesis-related proteins, European Journal of Plant Pathology, 103 (1997) 753-765.

[9] S. Ali, B.A. Ganai, A.N. Kamili, A.A. Bhat, Z.A. Mir, J.A. Bhat, A. Tyagi, S.T. Islam, M. Mushtaq, P. Yadav, S. Rawat, A. Grover, Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance, Microbiological Research, 212-213 (2018) 29-37.

[10] A. Edreva, Pathogenesis-related proteins: Research progress in the last 15 years, Gen Appl Plant Physiology, 31 (2005) 105-124.

[11] M. Sinha, R.P. Singh, G.S. Kushwaha, N. Iqbal, A. Singh, S. Kaushik, P. Kaur, S. Sharma, T.P. Singh, Current overview of allergens of plant pathogenesis related protein families, ScientificWorldJournal, 2014 (2014) 543195-543195.

[12] Hakim, A. Ullah, A. Hussain, M. Shaban, A.H. Khan, M. Alariqi, S. Gul, Z. Jun, S. Lin, J. Li, S. Jin, M.F.H. Munis, Osmotin: A plant defense tool against biotic and abiotic stresses, Plant Physiology and Biochemistry, 123 (2018) 149-159.

[13] J. Grenier, A. Asselin, Sorne Pathogenesis-Related Proteins Are Chitosanases with Lytic Activity Against Fungal Spores, MOLECULAR PLANT-MICROBE INTERACTIONS, 3 (1990) 401-407.

[14] T. Kawase, S. Yokokawa, A. Saito, T. Fujii, N. Nikaidou, K. Miyashita, T. Watanabe, Comparison of Enzymatic and Antifungal Properties between Family 18 and 19 Chitinases from S. coelicolor A3(2), Bioscience, Biotechnology, and Biochemistry, 70 (2006) 988-998.

[15] K. Min, S.C. Ha, P.M. Hasegawa, R.A. Bressan, D.-J. Yun, K.K. Kim, Crystal structure of osmotin, a plant antifungal protein, Proteins: Structure, Function, and Bioinformatics, 54 (2004) 170-173.
[16] D.N. Patil, M. Datta, A. Dev, S. Dhindwal, N. Singh, P. Dasauni, S. Kundu, A.K. Sharma, S. Tomar, P. Kumar, Structural Investigation of a Novel N-Acetyl Glucosamine Binding Chi-Lectin Which Reveals Evolutionary Relationship with Class III Chitinases, PLoS ONE, 8 (2013) e63779.
[17] J.-J. Liu, R. Sturrock, A.M. Ekramoddoullah, The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function, Plant Cell Reports, 29 (2010) 419-436.
[18] T. Pusztahelyi, Chitin and chitin-related compounds in plant-fungal interactions, Mycology, 9 (2018) 189-201.
[19] A. Sánchez-Vallet, J.R. Mesters, B.P.H.J. Thomma, The battle for chitin recognition in plant-microbe interactions, FEMS Microbiology Reviews, 39 (2015) 171-183.
[20] M. Malerba, R. Cerana, Recent Applications of Chitin- and Chitosan-Based Polymers in Plants, Polymers, 11 (2019).
[21] M.S. Riaz Rajoka, L. Zhao, H.M. Mehwish, Y. Wu, S. Mahmood, Chitosan and its derivatives: synthesis, biotechnological applications, and future challenges, Applied Microbiology and Biotechnology, 103 (2019) 1557-1571.
[22] D.S. Gomes, M.A. Lopes, S.P. Menezes, L.F. Ribeiro, C.V. Dias, B.S. Andrade, R.M. de Jesus, A.B.L. Pires, A. Goes-Neto, F. Micheli, Mycelial development preceding basidioma formation in Moniliophthora perniciosa is associated to chitin, sugar and nutrient metabolism alterations involving autophagy, Fungal Genetics and Biology, 86 (2016) 33-46.
[23] M.A. Lopes, D.S. Gomes, M.G. Koblitz, C.P. Pirovani, J.C. Cascardo, A. Goes-Neto, F. Micheli, Use of response surface methodology to examine chitinase regulation in the basidiomycete *Moniliophthora perniciosa*, Mycol Res, 112 (2008) 399-406.
[24] C.S. Souza, B.M. Oliveira, G.G. Costa, A. Schriefer, A. Selbach-Schnadelbach, A.P. Uetanabaro, C.P. Pirovani, G.A. Pereira, A.G. Taranto, J.C. Cascardo, A. Goes-Neto, Identification and characterization of a class III chitin synthase gene of *Moniliophthora perniciosa*, the fungus that causes witches' broom disease of cacao, J Microbiol, 47 (2009) 431-440.
[25] D. Penman, G. Britton, K. Hardwick, H.A. Collin, S. Isaac, Chitin as a measure of biomass of Crinipellis perniciosa, causal agent of witches' broom disease of Theobroma cacao, Mycological Research, 104 (2000) 671-675.
[26] L. Ferraz Dos Santos, R. Moreira Fregapani, L.L. Falcão, R.C. Togawa, M.M.d.C. Costa, U.V. Lopes, K. Peres Gramacho, R.M. Alves, F. Micheli, L.H. Marcellino, First Microsatellite Markers Developed from Cupuassu ESTs: Application in Diversity Analysis and Cross-Species Transferability to Cacao, PLoS ONE, 11 (2016) e0151074-e0151074.
[27] L. Ferraz dos Santos, R.J. Santana Silva, D. Oliveira Jordão do Amaral, M.F. Barbosa de Paula, L. Ludke Falcão, T. Legavre, R.M. Alves, L.H. Marcellino, F. Micheli, Selection of Reference Genes for Expression Study in Pulp and Seeds of Theobroma grandiflorum (Willd. ex Spreng.) Schum, PLoS ONE, 11 (2016) e0160646.
[28] L.L. Falcao, J.O.S. Werneck, R.C. Togawa, M.M.d.C. Costa, P. Grynberg, O.B.d. Silva Junior, R.M. Alves, P.S.B. Albuquerque, L.H. Marcellino, Analyses of cupuassu (*Theobroma grandiflorum*) transcriptome during interaction with *Moniliophthora perniciosa*, the causal agent of Witches' Broom disease., in: S.-E. CPATU (Ed.) 7th BRAZILIAN BIOTECHNOLOGY CONGRESS and 2nd BIOTECHNOLOGY IBERO-AMERICAN CONGRESS, Brasília, DF, Brazil, 2018.
[29] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat Meth, 8 (2011) 785-786.
[30] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, B. A., Protein Identification and Analysis Tools on the ExPASy Server, in: J.M.W. (ed) (Ed.) The Proteomics Protocols Handbook, Humana Press 2005, pp. 571-607.
[31] S. Briesemeister, J. Rahnenführer, O. Kohlbacher, YLoc--an interpretable web server for predicting subcellular localization, Nucleic Acids Research, 38 (2010) W497-W502.
[32] X. Argout, J. Salse, J.-M. Aury, M.J. Guiltinan, G. Droc, J. Gouzy, M. Allegre, C. Chaparro, T. Legavre, S.N. Maximova, M. Abrouk, F. Murat, O. Fouet, J. Poulain, M. Ruiz, Y. Roguet, M. Rodier-

520 Goud, J.F. Barbosa-Neto, F. Sabot, D. Kudrna, J.S.S. Ammiraju, S.C. Schuster, J.E. Carlson, E. Sallet, T.
521 Schiex, A. Dievart, M. Kramer, L. Gelley, Z. Shi, A. Bérard, C. Viot, M. Boccara, A.M. Risterucci, V.
522 Guignon, X. Sabau, M.J. Axtell, Z. Ma, Y. Zhang, S. Brown, M. Bourge, W. Golser, X. Song, D. Clement,
523 R. Rivallan, M. Tahi, J.M. Akaza, B. Pitollat, K. Gramacho, A. D'Hont, D. Brunel, D. Infante, I. Kebe, P.
524 Costet, R. Wing, W.R. McCombie, E. Guiderdoni, F. Quetier, O. Panaud, P. Wincker, S. Bocs, C.
525 Lanaud, The genome of Theobroma cacao, Nature Genetics, 43 (2011) 101-108.
526 [33] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics
527 Analysis across Computing Platforms, Molecular Biology and Evolution, 35 (2018) 1547-1549.
528 [34] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based
529 environment for protein structure homology modelling, Bioinformatics, 22 (2006) 195-201.
530 [35] F. MELO, d. Melo, e. melo, g. melo, *ANOLEA*: a www server to assess protein structures, *Intell.*
531 *Sys. Mol. Biol*., 97 (1997) 110-113.
532 [36] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the
533 stereochemical quality of protein structures, Journal of Applied Crystallography, 26 (1993) 283-291.
534 [37] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin,
535 I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X
536 version 2.0, Bioinformatics, 23 (2007) 2947-2948.
537 [38] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson,
538 AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, Journal of
539 Computational Chemistry, 30 (2009) 2785-2791.
540 [39] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new
541 scoring function, efficient optimization and multithreading, Journal of Computational Chemistry, 31
542 (2010) 455-461.
543 [40] W.L. DeLano, The PyMOL Molecular Graphics System, in, 2002.
544 [41] E.D. Cruz, R.M. Alvez, R.L. Benchimol, Avaliação de clones de cupuacuzeiro (Theobroma
545 grandiflorum (Willd. ex Spreng) K. Schumm.) quanto a tolerância a vassoura-de-bruxa (Crinipellis
546 perniciosa (Stahel) Singer). in: Embrapa Amazônia Oriental. Comunicado técnico, 28, Embrapa
547 CPATU, Belém, PA, Brazil, 2000, pp. 4 p.
548 [42] S. Surujdeo-Maharaj, P. Umaharan, D.R. Butler, T.N. Sreenivasan, An optimized screening
549 method for identifying levels of resistance to *Crinipellis perniciosa* in cocoa (*Theobroma cacao*), Plant
550 Pathol, 52 (2003) 464-475.
551 [43] G. Frias, L. Purdy, R. Schmidt, An inoculation method for evaluating resistance of cacao to
552 *Crinipellis perniciosa*, Plant Disease, 79 (1995) 787-791.
553 [44] A. Untergasser, H. Nijveen, X. Rao, T. Bisseling, R. Geurts, J.A.M. Leunissen, Primer3Plus, an
554 enhanced web interface to Primer3, Nucleic Acids Research, 35 (2007) W71-W74.
555 [45] B. Thornton, C. Basu, Real-time PCR (qPCR) primer design using free online software,
556 Biochemistry and Molecular Biology Education, 39 (2011) 145-154.
557 [46] S. Zhao, R.D. Fernald, Comprehensive algorithm for quantitative real-time polymerase chain
558 reaction, J Comput Biol, 12 (2005) 1047-1064.
559 [47] C.L. Andersen, J.L. Jensen, T.F. Ørntoft, Normalization of Real-Time Quantitative Reverse
560 Transcription-PCR data: a model-based variance estimation approach to identify genes suited for
561 normalization, applied to bladder and colon cancer data sets, Cancer Research, 64 (2004) 5245-5250.
562 [48] M.W. Pfaffl, G.W. Horgan, L. Dempfle, Relative expression software tool (REST) for group-wise
563 comparison and statistical analysis of relative expression results in real-time PCR, Nucleic Acids
564 Research, 30 (2002) e36-e36.
565 [49] N. Takahashi-Ando, M. Inaba, S. Ohsato, T. Igawa, R. Usami, M. Kimura, Identification of multiple
566 highly similar XIP-type xylanase inhibitor genes in hexaploid wheat, Biochemical and Biophysical
567 Research Communications, 360 (2007) 880-884.
568 [50] S. Martínez-Caballero, P. Cano-Sánchez, I. Mares-Mejía, A.G. Díaz-Sánchez, M.L. Macías-
569 Rubalcava, J.A. Hermoso, A. Rodríguez-Romero, Comparative study of two GH19 chitinase-like
570 proteins from Hevea brasiliensis, one exhibiting a novel carbohydrate-binding domain, The FEBS
571 Journal, 281 (2014) 4535-4554.

572 [51] T. Ohnuma, N. Umemoto, T. Nagata, S. Shinya, T. Numata, T. Taira, T. Fukamizo, Crystal structure
573 of a “loopless” GH19 chitinase in complex with chitin tetrasaccharide spanning the catalytic center,
574 Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1844 (2014) 793-802.
575 [52] F. Fukamizo, Chitinolytic Enzymes: Catalysis, Substrate Binding, and their Application, Current
576 Protein & Peptide Science, 1 (2000) 105-124.
577 [53] T. Masuda, G. Zhao, B. Mikami, Crystal structure of class III chitinase from pomegranate provides
578 the insight into its metal storage capacity, Bioscience, Biotechnology, and Biochemistry, 79 (2015) 45-
579 50.
580 [54] H. Koiwa, H. Kato, T. Nakatsu, J.i. Oda, Y. Yamada, F. Sato, Crystal structure of tobacco PR-5d
581 protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like
582 proteins11Edited by R. Huber, Journal of Molecular Biology, 286 (1999) 1137-1145.
583 [55] M.M. Chaudet, T.A. Naumann, N.P.J. Price, D.R. Rose, Crystallographic structure of ChitA, a
584 glycoside hydrolase family 19, plant class IV chitinase from Zea mays, Protein Science, 23 (2014) 586-
585 593.
586 [56] N.A. Udaya Prakash, M. Jayanthi, R. Sabarinathan, P. Kangueane, L. Mathew, K. Sekar, Evolution,
587 Homology Conservation, and Identification of Unique Sequence Signatures in GH19 Family
588 Chitinases, Journal of Molecular Evolution, 70 (2010) 466-478.
589 [57] A.S. Fister, L.C. Mejia, Y. Zhang, E.A. Herre, S.N. Maximova, M.J. Guiltinan, Theobroma cacao L.
590 pathogenesis-related gene tandem array members show diverse expression dynamics in response to
591 pathogen colonization, BMC Genomics, 17 (2016) 363-363.
592 [58] K. Koizumi, R. Yokoyama, K. Nishitani, Mechanical load induces upregulation of transcripts for a
593 set of genes implicated in secondary wall formation in the supporting tissue of Arabidopsis thaliana,
594 Journal of Plant Research, 122 (2009) 651.
595 [59] Y. Kwon, S.-H. Kim, M.-S. Jung, M.-S. Kim, J.-E. Oh, H.-W. Ju, K.-i. Kim, E. Vierling, H. Lee, S.-W.
596 Hong, Arabidopsis hot2 encodes an endochitinase-like protein that is essential for tolerance to heat,
597 salt and drought stresses, The Plant Journal, 49 (2007) 184-193.
598 [60] C. Sánchez-Rodríguez, S. Bauer, K. Hématy, F. Saxe, A.B. Ibáñez, V. Vodermaier, C. Konlechner, A.
599 Sampathkumar, M. Rüggeberg, E. Aichinger, L. Neumetzler, I. Burgert, C. Somerville, M.-T. Hauser, S.
600 Persson, Chitinase-like1/pom-pom1 and its homolog CTL2 are glucan-interacting proteins important
601 for cellulose biosynthesis in Arabidopsis, The Plant Cell, 24 (2012) 589-607.
602 [61] R.F. Ditt, K.F. Kerr, P. de Figueiredo, J. Delrow, L. Comai, E.W. Nester, The Arabidopsis thaliana
603 Transcriptome in Response to Agrobacterium tumefaciens, Molecular Plant-Microbe Interactions®,
604 19 (2006) 665-681.
605 [62] Z. Zheng, S. Qamar, Z. Chen, T. Mengiste, Arabidopsis WRKY33 transcription factor is required for
606 resistance to necrotrophic fungal pathogens, Plant Journal, 48 (2006).
607 [63] Y. Takenaka, S. Nakano, M. Tamoi, S. Sakuda, T. Fukamizo, Chitinase Gene Expression in
608 Response to Environmental Stresses in Arabidopsis thaliana: Chitinase Inhibitor Allosamidin Enhances
609 Stress Tolerance, Bioscience, Biotechnology, and Biochemistry, 73 (2009) 1066-1071.
610 [64] P.J.P.L. Teixeira, D.P.d.T. Thomazella, O. Reis, P.F.V. do Prado, M.C.S. do Rio, G.L. Fiorin, J. José,
611 G.G.L. Costa, V.A. Negri, J.M.C. Mondego, P. Mieczkowski, G.A.G. Pereira, High-resolution transcript
612 profiling of the atypical biotrophic interaction between *Theobroma cacao* and the fungal pathogen
613 *Moniliophthora perniciosa*, The Plant Cell, 26 (2014) 4245-4269.
614 [65] K. Sena, L. Alemanno, K.P. Gramacho, The infection process of *Moniliophthora perniciosa* in
615 cacao, Plant Pathology, (2014) 1272–1281.
616 [66] D. Prasath, I. El-Sharkawy, S. Sherif, K.S. Tiwary, S. Jayasankar, Cloning and characterization of
617 PR5 gene from Curcuma amada and Zingiber officinale in response to Ralstonia solanacearum
618 infection, Plant Cell Reports, 30 (2011) 1799.
619 [67] D. PRASATH, A. BALAGOPAL, V. MAHANTESH, O.B. ROSANA, S. JAYANSANKAR, M. ANANDARAJ,
620 Comparative study of pathogenesis-related protein 5 (PR5) of different Zingiberaceae species., Indian
621 Journal of Biotechnology, 13 (2014) 178-185.

[68] A. El-kereamy, I. El-sharkawy, R. Ramamoorthy, A. Taheri, D. Errampalli, P. Kumar, S. Jayasankar, Prunus domestica pathogenesis-related protein-5 activates the defense response pathway and enhances the resistance to fungal infection, PLoS ONE, 6 (2011) e17973-e17973.
[69] J. Grenier, C. Potvin, J. Trudel, A. Asselin, Some thaumatin-like proteins hydrolyse polymeric β-1,3-glucans, The Plant Journal, 19 (1999) 473-480.
[70] C.S. Barbosa, R.R.d. Fonseca, T.M. Batista, M.A. Barreto, C.S. Argolo, M.R.d. Carvalho, D.O.J.d. Amaral, E.M.d.A. Silva, E. Arévalo-Gardini, K.S. Hidalgo, G.R. Franco, C.P. Pirovani, F. Micheli, K.P. Gramacho, Genome sequence and effectorome of Moniliophthora perniciosa and Moniliophthora roreri subpopulations, BMC Genomics, 19 (2018) 509-509.
[71] R.M. Alves, M.D.V. Resende, B.S. Bandeira, T.M. Pinheiro, D.C.R. Farias, Evolução da vassoura-de-bruxa e avaliação da resistência em progênies de cupuaçuzeiro, Revista Brasileira de Fruticultura, 31 (2009) 1022-1032.
[72] H. Böhm, I. Albert, L. Fan, A. Reinhard, T. Nürnberger, Immune receptor complexes at the plant cell surface, Current Opinion in Plant Biology, 20 (2014) 47-54.
[73] H. Kaku, Y. Nishizawa, N. Ishii-Minami, C. Akimoto-Tomiyama, N. Dohmae, K. Takio, E. Minami, N. Shibuya, Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor, Proceedings of the National Academy of Sciences of the United States of America, 103 (2006) 11086-11091.

641 **Tables**

642

643 **Table 1.** TgPR gene and protein sequence characteristics.

| Name | Gene data | | | Protein data | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Identification* | mRNA * (bp) | ORF size (bp) | Protein size (aa) | Signal peptide (aa) | Molecular weight (kDa) | pI | Subcellular localization (probability) | Phosphorylation sites | Glycosylation sites |
| TgPR3 | C106 | 1239 | 966 | 322 | 26 | 35.5 | 6.64 | Extracellular (86.48%) | $S_{37}$ $S_{41}$ $S_{114}$ $S_{123}$ $S_{193}$ $T_{91}$ $T_{99}$ $T_{185}$ $T_{205}$ $T_{219}$ $T_{240}$ $Y_{62}$ $Y_{119}$ $Y_{124}$ $Y_{126}$ $Y_{145}$ $Y_{167}$ | $N_{20}$ $N_{171}$ $N_{207}$ |
| TgPR5 | C68 | 950 | 672 | 224 | 23 | 24.2 | 6.68 | Extracellular (99.7%) | $S_{118}$ $S_{139}$ $S_{161}$ $S_{176}$ $T_{143}$ $T_{256}$ $Y_{137}$ $Y_{239}$ | $N_{163}$ |
| TgPR8 | C356 | 1111 | 927 | 309 | 29 | 33.3 | 4.27 | Extracellular (52.2%) Vacuolar (34.5%) | $S_{77}$ $S_{95}$ $S_{147}$ $S_{266}$ $T_{15}$ $T_{41}$ $T_{88}$ $Y_{66}$ $Y_{99}$ $Y_{134}$ $Y_{149}$ $Y_{162}$ $Y_{227}$ | $N_{145}$ |

644 * Data from RNA-seq as described by [28].

645 **Figure legends**

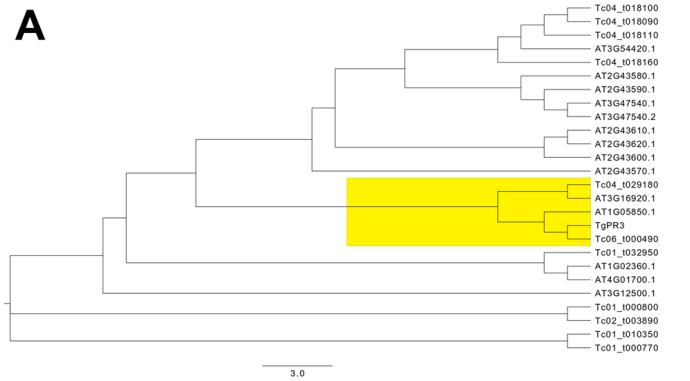
646 **Figure 1.** Phylogenetic tree obtained with the *TgPR* nucleotide sequences and their
647 homologues from *T. cacao* and *A. thaliana*. **A.** Phylogenic tree from *TgPR3*. **B.**
648 Phylogenic tree from *TgPR5*. **C.** Phylogenic tree from *TgPR8*. Sequence names
649 beginning with Tc and AT correspond to *T. cacao* and *A. thaliana*, respectively. The
650 yellow square highlights the TgPR's closest phylogenetic branches.

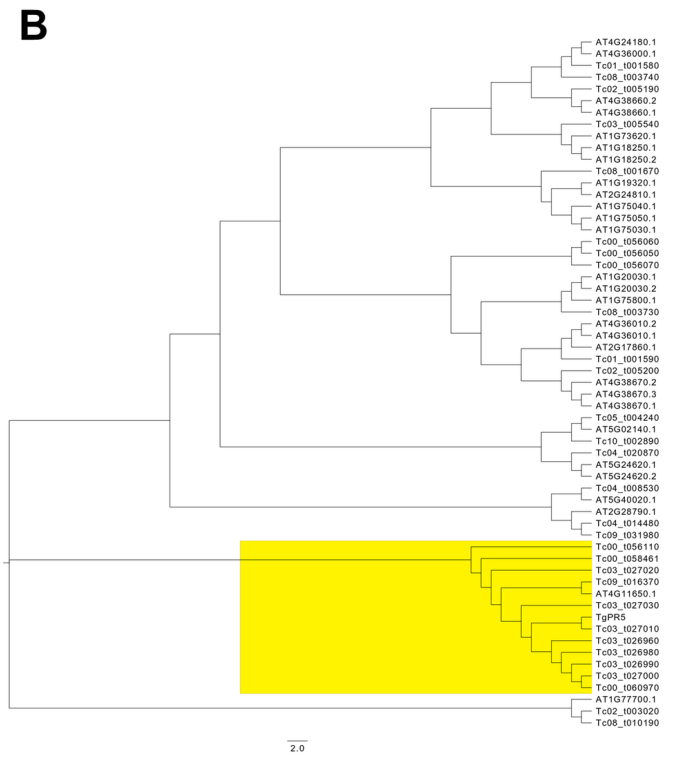
651 **Figure 2.** The 3D structure of TgPR proteins. **A.** TgPR3's 3D structure, with the amino
652 acid residues Lys70 Glu92, and Tyr125 of the active site highlighted. **B.** TgPR5's 3D
653 structure with the amino acid residues Arg43, Glu83, Asp96 and Asp101 of the active
654 site highlighted. **C.** TgPR8's 3D structure with the amino acid residues Tyr6, Phe32,
655 Asp125, Glu127, Gln155, Gln178, Tyr180 and Trp253 of the active site highlighted.

656 **Figure 3.** Representations (2D and 3D) of docking between TgPRs and chitin/chitosan.

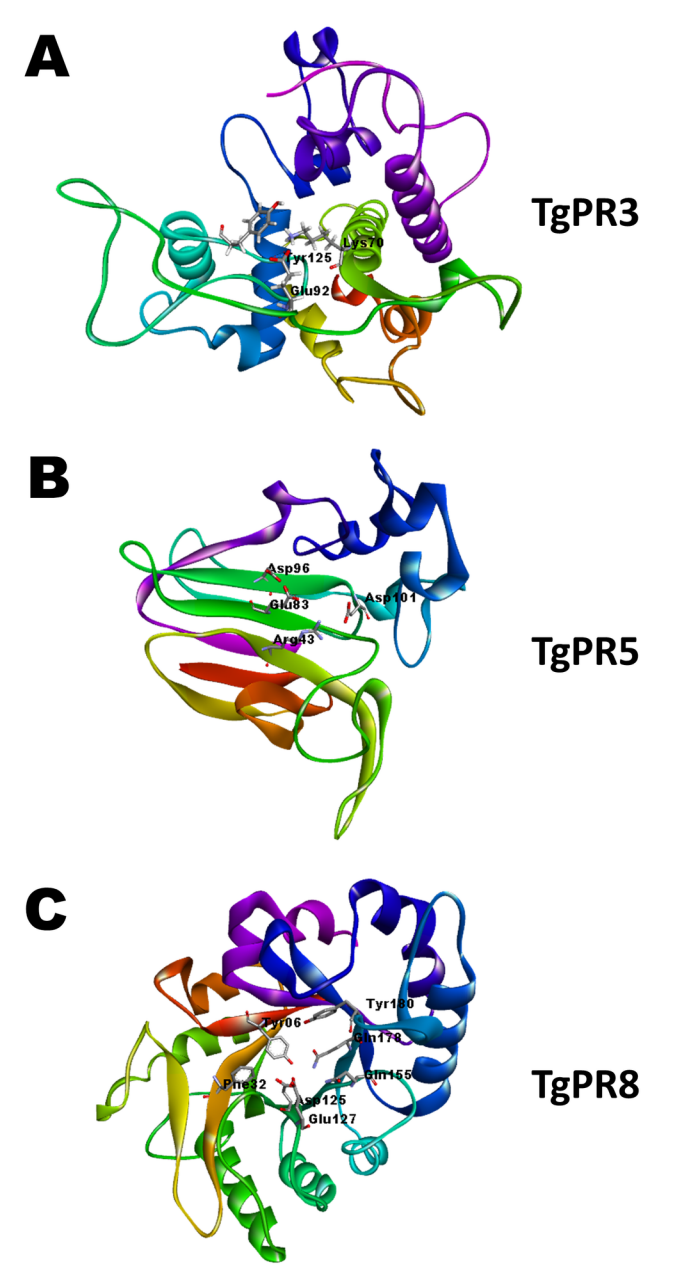
657 **Figure 4.** Relative expression of TgPR genes in resistant (R) and susceptible (S)
658 cupuassu genotypes 8, 24, 48 and 72 hours after inoculation (hai) with *Moniliophthora*
659 *perniciosa*. **A.** Full-scale representation. **B.** Reduced scale for the best visualization of
660 lower relative expression.

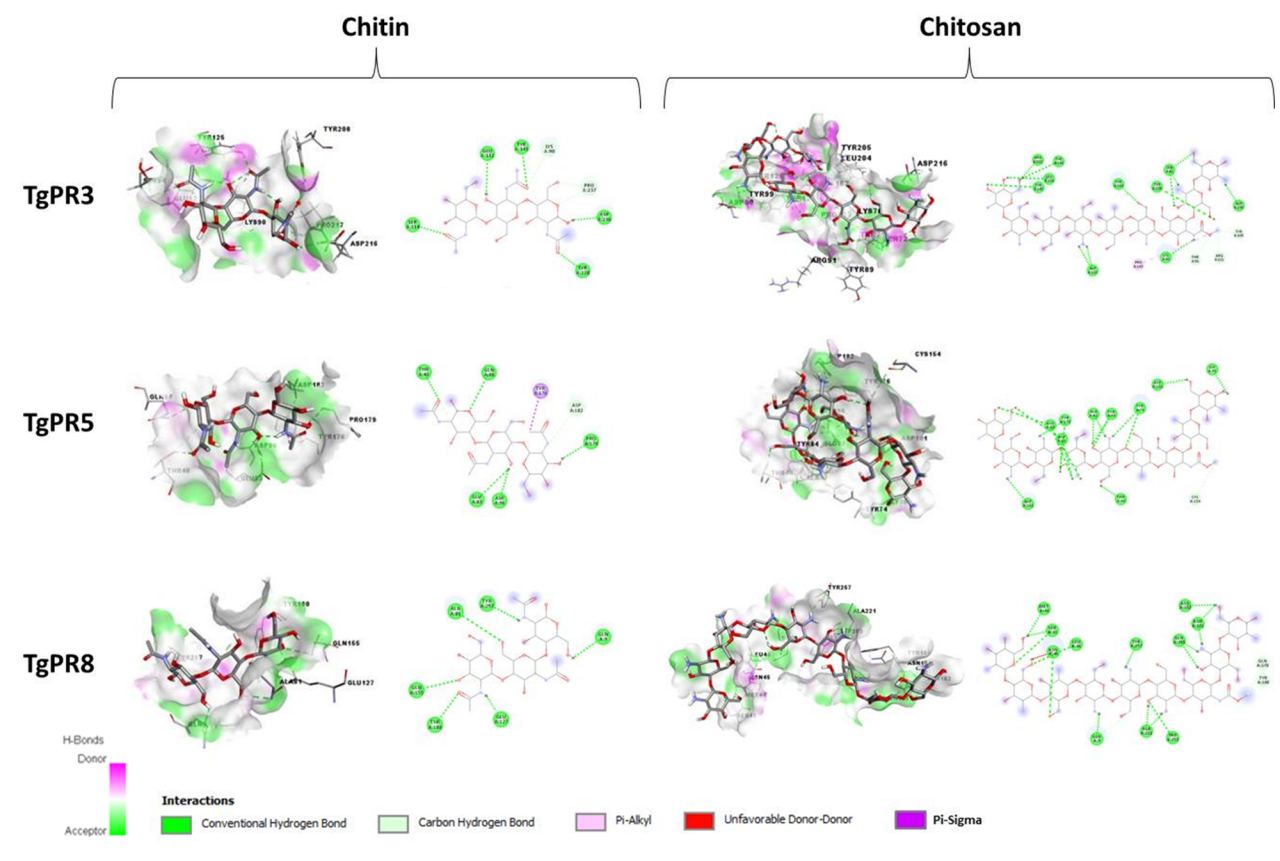
661 **Figure 5.** Putative scheme of the early steps (6–72 hai) of the interactions between *T.*
662 *grandiflorum* plants and the hemibiotrophic fungus *M. perniciosa*. **A.** Resistant
663 interaction. **B.** Susceptible interaction. PTI: PAMP-triggered immunity; ETI: effector-
664 triggered immunity; PRR: pattern-recognition receptors; PR: pathogenesis-related
665 protein; PCD: programmed cell death.

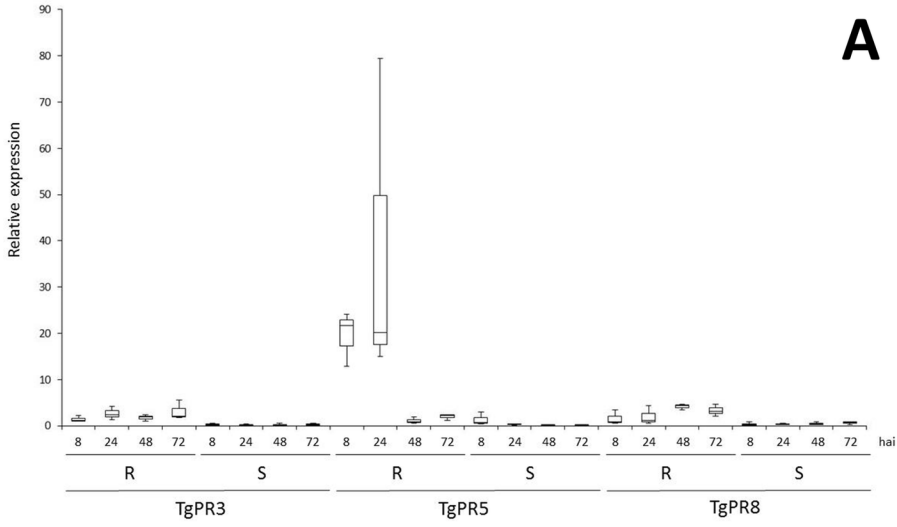

A
Tc04_t018100
Tc04_t018090
Tc04_t018110
AT3G54420.1
Tc04_t018160
AT2G43580.1
AT2G43590.1
AT3G47540.1
AT3G47540.2
AT2G43610.1
AT2G43620.1
AT2G43600.1
AT2G43570.1
Tc04_t029180
AT3G16920.1
AT1G05850.1
TgPR3
Tc06_t000490
Tc01_t032950
AT1G02360.1
AT4G01700.1
AT3G12500.1
Tc01_t000800
Tc02_t003890
Tc01_t010350
Tc01_t000770
3.0

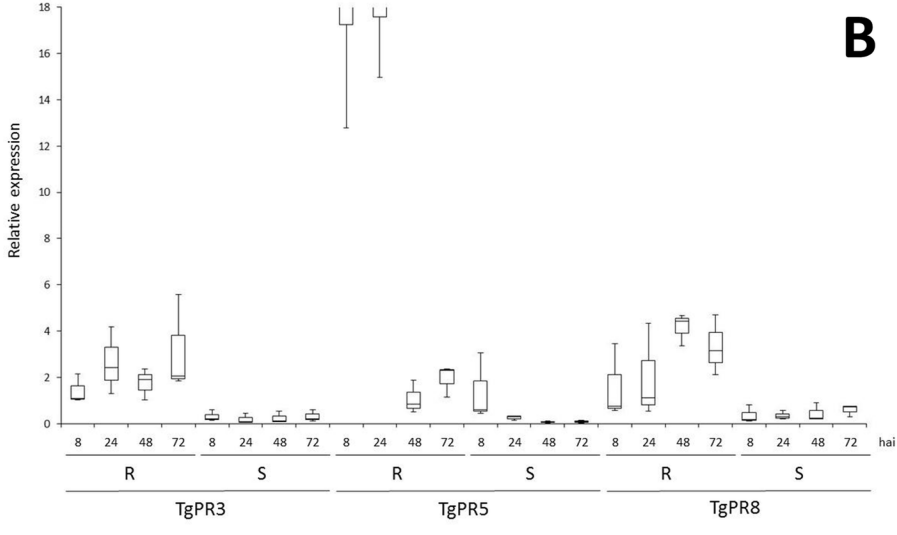

B
AT4G24180.1
AT4G36000.1
Tc01_t001580
Tc08_t003740
Tc02_t005190
AT4G38660.2
AT4G38660.1
Tc03_t005540
AT1G73620.1
AT1G18250.1
AT1G18250.2
Tc08_t001670
AT1G19320.1
AT2G24810.1
AT1G75040.1
AT1G75050.1
AT1G75030.1
Tc00_t056060
Tc00_t056050
Tc00_t056070
AT1G20030.1
AT1G20030.2
AT1G75800.1
Tc08_t003730
AT4G36010.2
AT4G36010.1
AT2G17860.1
Tc01_t001590
Tc02_t005200
AT4G38670.2
AT4G38670.3
AT4G38670.1
Tc05_t004240
AT5G02140.1
Tc10_t002890
Tc04_t020870
AT5G24620.1
AT5G24620.2
Tc04_t008530
AT5G40020.1
AT2G28790.1
Tc04_t014480
Tc09_t031980
Tc00_t056110
Tc00_t058461
Tc03_t027020
Tc09_t016370
AT4G11650.1
Tc03_t027030
TgPR5
Tc03_t027010
Tc03_t026960
Tc03_t026980
Tc03_t026990
Tc03_t027000
Tc00_t060970
AT1G77700.1
Tc02_t003020
Tc08_t010190
2.0

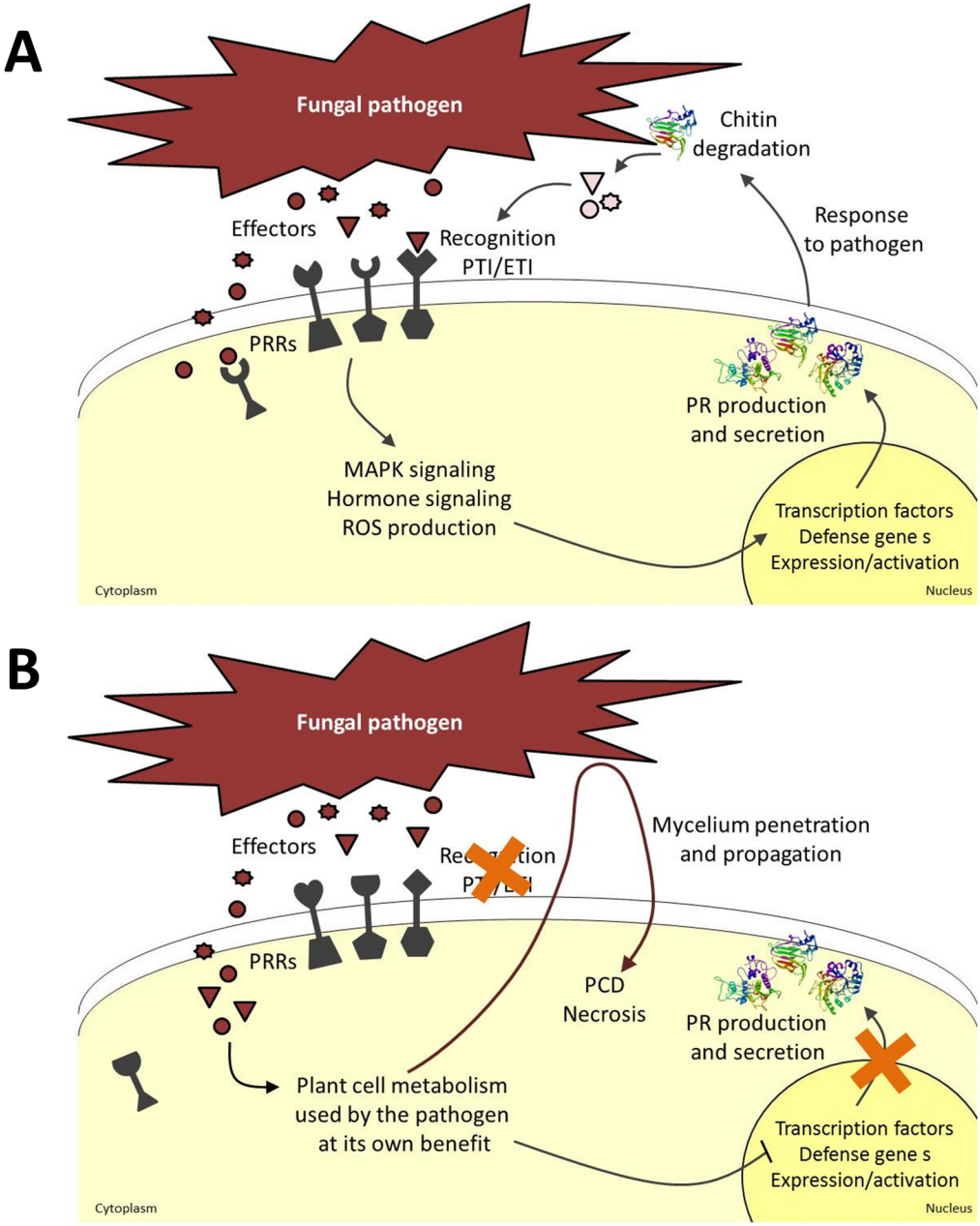

C
TgPR8
Tc01_t032120
Tc10_t015330
Tc10_t015260
Tc03_t017780
Tc03_t017790
Tc03_t017760
AT5G24090.1
Tc04_t001620
Tc01_t035140
Tc00_t024510
Tc01_t035150
Tc01_t035160
Tc00_t027770
Tc01_t035050
Tc01_t035060
0.7


A
Lys70
Tyr125
Glu92
TgPR3
B
Asp96
Glu83
Asp101
Arg43
TgPR5
C
Tyr130
Tyr106
Glu179
Phe32
Glu155
Asp125
Glu127
TgPR8


Chitin
Chitosan
TgPR3
TgPR5
TgPR8
H-Bonds
Donor
Acceptor
Interactions
Conventional Hydrogen Bond
Carbon Hydrogen Bond
Pi-Alkyl
Unfavorable Donor-Donor
Pi-Sigma


A
Relative expression
90
80
70
60
50
40
30
20
10
0
8 24 48 72 8 24 48 72 8 24 48 72 8 24 48 72 8 24 48 72 8 24 48 72
hai
R S R S R S
TgPR3 TgPR5 TgPR8


B
Relative expression
18
16
14
12
10
8
6
4
2
0
8 24 48 72 8 24 48 72 8 24 48 72 8 24 48 72 8 24 48 72 8 24 48 72
hai
R S R S R S
TgPR3 TgPR5 TgPR8


A
Fungal pathogen
Chitin degradation
Effectors
Recognition PTI/ETI
Response to pathogen
PRRs
PR production and secretion
MAPK signaling
Hormone signaling
ROS production
Transcription factors
Defense gene s
Expression/activation
Cytoplasm
Nucleus
B
Fungal pathogen
Mycelium penetration and propagation
Effectors
PRRs
PCD
Necrosis
PR production and secretion
Plant cell metabolism used by the pathogen at its own benefit
Transcription factors
Defense gene s
Expression/activation
Cytoplasm
Nucleus