Natural variation in insect egg-induced cell death uncovers a role for L-type LECTIN RECEPTOR KINASE-I.1 in Arabidopsis

Raphaël Groux[1], Caroline Gouhier-Darimont[1], Envel Kerdaffrec[2] and Philippe Reymond[1]

[1]Department of Plant Molecular Biology, University of Lausanne, CH-1015 Lausanne, Switzerland; [2]Department of Biology, University of Fribourg, CH-1700 Fribourg, Switzerland

Author for correspondence:
Philippe Reymond
Tel: +41 21 692 41 90
Email: philippe.reymond@unil.ch

Abstract

In *Arabidopsis thaliana*, a hypersensitive-like response (HR-like) is triggered underneath the eggs of the large white butterfly *Pieris brassicae*, and this response is dependent on salicylic acid (SA) accumulation and signaling. Previous reports indicate that the clade I L-type lectin receptor kinase LecRK-I.8 is involved in early steps of egg recognition. A genome-wide association study (GWAS) was used to better characterize the genetic structure of HR-like and discover loci that contribute to this response. We report here the identification of LecRK-I.1, a close homolog of LecRK-I.8, and show that two main haplotypes that explain part of the variation in HR-like segregate amongst natural Arabidopsis accessions. In addition, signatures of balancing selection at this locus suggest that it may be ecologically important. Disruption of LecRK-I.1 resulted in decreased HR-like while SA-dependent signaling was unaffected, suggesting that this protein is specifically involved in the induction of cell death. Furthermore, we provide evidence that LecRK-I.1 functions in the same signaling pathway as LecRK-I.8. Altogether, our results show that the response to eggs of *P. brassicae* is controlled by LecRKs that operate at various steps of the signaling pathway.

Key words: Arabidopsis, genome-wide association study, hypersensitive-like response, insect eggs, LecRK-I.1, *Pieris brassicae*

38 **Introduction**

39 Although eggs of herbivorous insects deposited on plant leaves are immobile and inert
40 structures, they represent a future threat when hatching larvae start to feed. Plants are able to
41 detect the presence of eggs and respond by triggering direct and indirect defenses (Reymond,
42 2013; Hilker & Fatouros, 2015). Direct defenses such as tissue growth and localized cell death
43 lead to reduced egg hatching, egg crushing by surrounding tissues, or egg desiccation and drop
44 off (Shapiro & DeVay, 1987; Doss *et al.,* 2000; Garza *et al.,* 2001; Desurmont & Weston, 2011;
45 Petzold-Maxwell *et al.,* 2011; Fatouros *et al.,* 2014). The production of ovicidal benzyl cyanide
46 at the oviposition site was reported in *Oryza sativa* (Suzuki *et al.,* 1996; Yamasaki *et al.,* 2003).
47 For indirect defenses, a release of volatiles or changes in leaf surface chemistry attract natural
48 egg predators (Hilker & Meiners, 2006; Fatouros *et al.,* 2012). While these defense mechanisms
49 impact egg mortality individually, some studies found that the co-induction of both direct and
50 indirect defense strategies synergistically impact egg survival (Fatouros *et al.,* 2014). Even
51 though such defenses can efficiently reduce insect pressure before damage occurs, one study
52 indicates that these traits may get lost during domestication (Tamiru *et al.,* 2015), similar to
53 what is observed for other defense-related traits (Chen *et al.,* 2015). Introgression of egg-killing
54 traits in cultivated varieties has been reported in *O. sativa* (Suzuki *et al.,* 1996; Yamasaki *et al.,*
55 2003; Yang *et al.,* 2014), however it is still a mostly unexploited strategy due to a lack of
56 mechanistic understanding of these responses at molecular and cellular levels (Reymond, 2013;
57 Fatouros *et al.,* 2016).

58 Plants from the Brassicales and Solanales orders respond to oviposition through the
59 induction of cell death (Petzold-Maxwell *et al.,* 2011; Fatouros *et al.,* 2014, 2016; Kalske *et al.,*
60 2014), a process that resembles the hypersensitive response (HR) triggered by certain adapted
61 pathogens. Based on this intriguing similarity, this type of insect egg-triggered response is
62 referred to as hypersensitive-like response (HR-like) (Reymond, 2013; Fatouros *et al.,* 2014).
63 Recent progress in the study of HR-like in response to insect eggs revealed that reactive oxygen
64 species (ROS) and salicylic acid (SA), two major signaling molecules regulating plant
65 immunity, accumulate at oviposition sites (Little *et al.,* 2007; Bruessow *et al.,* 2010; Geuss *et*
66 *al.,* 2017; Bittner *et al.,* 2017; Bonnet *et al.,* 2017). In *Arabidopsis thaliana*, cell death induction
67 upon crude egg extract (EE) treatment was reduced in mutants impaired in SA biosynthesis and
68 signaling (Gouhier-Darimont *et al.,* 2013). In *Brassica nigra*, only plants expressing HR-like
69 symptoms displayed elevated *PR1* transcript levels, a widely used SA marker gene, again
70 suggesting that SA accumulation is required for HR-like induction (Fatouros *et al.,* 2014).
71 These results indicate that the response to insect eggs is similar to pathogen-triggered immunity

72 (PTI), as demonstrated in Arabidopsis (Gouhier-Darimont *et al.,* 2013). However, knowledge
73 of elicitor-receptor pairs involved in plant response to eggs is lacking. A recent report identified
74 the Arabidopsis L-type LecRK-I.8 as an early signaling component of *Pieris brassicae* egg-
75 induced responses, thereby constituting a potential candidate for the perception of egg-derived
76 elicitors (Gouhier-Darimont *et al.,* 2019). Lectin-like receptor kinases (LecRKs) have been
77 implicated in a myriad of immune-related processes such as PTI (Takahashi *et al.,* 2007), chitin
78 perception (Miya *et al.,* 2007), e-(extracellular) ATP and $eNAD^+$/$eNADP^+$ perception (Choi *et*
79 *al.,* 2014; Wang *et al.,* 2019; Pham *et al.,* 2020), bacterial short chain 3OH-FA perception
80 (Kutschera *et al.,* 2019), and insect resistance (Gilardoni *et al.,* 2011; Liu *et al.,* 2015).
81 Although insect eggs trigger a PTI-like response in Arabidopsis, identification of early
82 and late components is still needed. Transcriptome changes after *P. brassicae* oviposition and
83 pathogen infection are similar but not identical, suggesting some level of specificity (Little *et*
84 *al.,* 2007). In support of this hypothesis, ROS production following *P. brassicae* EE treatment
85 was independent on the NADPH oxidases RBOHD/F (Gouhier-Darimont *et al.,* 2013), while
86 this pathway is crucial for pathogen-triggered ROS accumulation (Torres *et al.,* 2002; Morales
87 *et al.,* 2016). Moreover, a *lecrk-I.8* mutant plant did not show altered resistance to a bacterial
88 pathogen (Gouhier-Darimont *et al.,* 2019), suggesting that LecRK-I.8 may be specifically
89 involved in egg signaling responses. To get more molecular insight in insect-egg induced
90 response, we used the existing natural variation in EE-triggered HR-like in Arabidopsis and
91 performed a genome-wide association study (GWAS) using a set of 295 natural accessions. We
92 found a peak of association on chromosome 3 encompassing the clade I *L-type LecRK-I.*1, a
93 close homolog of *LecRK-I.8.* Subsequent experimental validation showed that LecRK-I.1 is
94 specifically involved in the regulation of cell death following insect egg perception, hence
95 providing a novel component of this immune response.
96
97 **Materials and Methods**
98
99 *Plant materials and growth conditions*
100 All experiments described here were carried out in *Arabidopsis thaliana*. Plants were grown in
101 growth chambers in short day conditions (10 h light, 22°C, 65% relative humidity, 100 µmol
102 m^{-2} s^{-1}) and were 4 to 5 weeks old at the time of treatment. Larvae, eggs and butterflies of the
103 Large White butterfly *Pieris brassicae* came from a population maintained on *Brassica*
104 *oleracea* in a greenhouse (Reymond *et al.,* 2000).

105 Accessions used for GWAS mapping were obtained from the NASC stock center and
106 are listed in Supporting Information Table S1. T-DNA insertion lines for *lecrk-I.1*
107 (SALK_052123), *lecrk-I.2* (SAIL_847_F07), *lecrk-I.3* (SALK_087804C), *lecrk-I.4*
108 (SALK_091901), *lecrk-I.5* (GABI_777H06), *lecrk-I.6* (GABI_353G10) and *lecrk-I.8*
109 (SALK_066416) were described previously (Gouhier-Darimont *et al.,* 2019). Primers for
110 genotyping insertion lines were designed with the SIGNAL T-DNA verification tool for all
111 lines used in this study.
112 A dual sgRNA CRISPR Cas9 approach (Pauwels *et al.,* 2018) was used to create a 33
113 kb deletion on chromosome 3 between *LecRK-I.1* (At3g45330) and *LecRK-I.6* (At3g45440),
114 generating the sextuple mutant *ccI.1-I.6*. sgRNAs specific for *LecRK-I.1* and *LecRK-I.6* were
115 selected using CRISPR-P (http://crispr.hzau.edu.cn/CRISPR2/). Complementary oligos with 4
116 bp overhangs were annealed and inserted in the Gateway ENTRY sgRNA shuttle vectors
117 (Pauwels *et al.,* 2018) pMR217 (L1-R5) for sgRNA-LecRK-I.1 and pMR218 (L5-L2) for
118 sgRNA-LecRK-I.6 via a cut-ligation reaction with BbsI (New England Biolabs) and T4 DNA
119 ligase (New England Biolabs), generating two sgRNA modules pMR217-sgRNA-LecRK-I.1
120 and pMR218-sgRNA-LecRK-I.6. Using a Gateway Multisite LR reaction (Thermofisher), the
121 two sgRNA modules were combined with pDE-Cas9Km (Pauwels *et al.,* 2018) to get the final
122 expression clone that was inserted in *Agrobacterium tumefaciens* strain GV3101 for
123 transformation of Arabidopsis Col-0 by floral dipping. Seeds were selected on ½ MS with 50
124 µg/ml kanamycin under continuous day conditions for 12 days. Presence of a deletion was
125 tested by PCR, using specific primers CC-LecRK-I.1-Rv and CC-LecRK-I.6-Fw flanking the
126 deletion and generating a 465 bp fragment (Figure S1).
127
128 *Genome-wide association study and haplotype analysis*
129 For GWAS analysis, a set of 295 accessions from the RegMap panel (Horton *et al.,* 2012) were
130 used (Table S1). Pools of 30 accessions were phenotyped every week over 2 days. For each
131 accession, 3 leaves from 3 to 6 plants were treated with egg-extract diluted 1:1 with deionized
132 water leading to a total of 9 to 18 treated leaves. Treated plants were left in the growth chamber
133 for an additional 5 days until phenotyping. After 5 days, treated leaves were removed with
134 forceps, symptoms were scored and pictures were taken as described below.
135 Mapping was performed locally using custom-made R scripts (R-Core Development
136 Team, 2005; Kerdaffrec *et al.*, 2016) with a linear mixed model that accounts for population
137 structure by computing a population-wide kinship matrix. Full imputed genotypes for 2029
138 lines (Togninalli *et al.*, 2018) were used and only SNPs with a minor-allele frequency (MAF)

139 > 0.05 were considered for analysis. After filtering, a total of 1'769'548 SNP were used for
140 mapping. To correct for genome-wide multiple testing, a Bonferroni corrected threshold of
141 significance was computed by dividing $\alpha = 0.05$ by the number of SNPs used in the analysis.
142 The percentage of variance explained by the top SNP (Chr. 3, 16633422) of *LecRK1-1* was
143 calculated using the residual sum of squares (RSS) extracted from the GWA model.
144 Haplotype analysis was based on the most significantly associated SNPs ($-\log_{10} P > 4$)
145 using data extracted from the genotype matrix. Presence/absence of premature stop codons in
146 sequenced accessions (Table S1) was recovered from the POLYMORPH1001 Variant browser
147 (https://tools.1001genomes.org/polymorph/). Both sets of data were used to explore phenotypic
148 data.
149 The population-wide cladogram was built by calculating the Euclidian distance between
150 accessions from the kinship matrix, and clustering was subsequently performed using the
151 "ward.d2" algorithm in R. Gene sequences in Fasta format from 125 accessions were obtained
152 from SALK 1001 genome browser (http://signal.salk.edu/atg1001/3.0/gebrowser.php) and
153 sequences were aligned using MAFFT for further analysis.
154 Sliding windows analyses of Tajima's D, Fu and Li's D / F on the LecRK-I.1 locus were
155 performed using DnaSP (version 6.12.03) with sliding windows of 200 variant sites and a step
156 size of 25 sites, and significance threshold was set according to a 95% confidence limit
157 published before (Tajima, 1989; Fu & Li, 1993). Homologous gene sequences from *AtLecRK-*
158 *I.1* in *Arabidopsis lyrata* were identified by BLAST, and the best hit was used as outgroup
159 sequence for Fu and Li's statistics. 50 kb sliding-window analysis of Tajima's D was computed
160 using VCF genotype file obtained from the 1001 Genome API
161 (https://tools.1001genomes.org/api/index.html). Tajima's D was computed on sliding windows
162 of 200 bp with a step size of 25 bp using custom-made R scripts.
163
164 *Broad sense heritability and variance analysis*
165 For each phenotype, broad sense heritability (H^2) was calculated using the following equations:
166 $H^2 = V_g/V_p$ and $V_p = V_g + V_e$, where V_g, V_p and V_e stand for genetic, phenotypic and
167 environmental variance respectively. Since accessions are homozygous, V_e was estimated as
168 the average intra-accession variance and V_p was the variance of the phenotype in the population.
169
170 *Oviposition and treatment with EE*
171 For experiments with natural oviposition, plants were placed in a tent containing approximately
172 20 *P. brassicae* butterflies for a maximum of 8 h. Afterwards, plants were kept in a growth

173 chamber in plastic boxes until hatching of the eggs. Control plants were kept in the same
174 conditions without butterflies.
175 *P. brassicae* eggs were collected and crushed with a pestle in Eppendorf tubes. After
176 centrifugation (15'000 g, 3 min), the supernatant ('egg extract', EE) was collected and stored at
177 -20°C. Plants were 4-5 weeks old at the time of treatment. For each plant, each of two leaves
178 were treated with 2 µl of EE. This amount corresponds to one egg batch of ca. 20 eggs. A total
179 of four plants were used for each experiment. After the appropriate time, EE was gently
180 removed with a scalpel blade and treated leaves were stored in liquid nitrogen. Untreated plants
181 were used as controls.
182 For GWAS, a large amount of EE was prepared as described and aliquots were stored
183 under N_2 at -80°C in order to ensure homogenous treatments during the entire experiment.
184
185 *Symptom scoring*
186 Symptoms were scored visually from the adaxial side of the leaves and were classified into the
187 following categories: no symptom (leaf treated area is still lush green, score = 0), small
188 chlorosis (<50% of the treated area, score = 1), large chlorosis (> 50%, score = 2), small necrosis
189 (brown spots or transparent membrane on < 50% of the treated area, score = 3), large necrosis
190 (> 50 %, score = 4) and spreading necrosis (necrosis not confined to the treated area, score =
191 5).
192
193 *Histochemical staining*
194 For visualization of cell death, EE was gently removed and leaves were submerged in
195 lactophenol trypan blue solution (5 ml of lactic acid, 10 ml of 50% glycerol, 1 mg of trypan
196 blue (Sigma), and 5 ml of phenol) at 28°C for 2–3 h. After staining, leaves were destained for
197 10 min in boiling 95% ethanol. Microscope images were saved as TIFF files and processed for
198 densitometric quantification with ImageJ (version 1.48).
199
200 *Salicylic acid quantifications*
201 SA quantifications were performed using the bacterial biosensor *Acinetobacter sp*. ADPWH
202 (Huang *et al.,* 2005; Huang *et al.,* 2006) according to (DeFraia *et al.,* 2008; Zvereva *et al.,*
203 2016). Briefly, 6 leaf discs (0.7 cm, ~20 mg) were ground in liquid nitrogen and extracted in
204 0.1M sodium acetate buffer (pH 5.6). Extracts were then centrifuged at 4°C for 15 min at 16'000
205 g. 50 µL of extract were incubated with 5 µL of β-Glucosidase from almonds (0.5 U/µl in
206 acetate buffer, Sigma-Aldrich) during 90 min at 37°C to release SA from SA-glucoside (SAG).

207 20 µL of extract was then mixed with 60 µL of LB and 50 µL of a culture of *Acinetobacter* sp.
208 ADPWH_lux (OD_{600} = 0.4), and incubated for 1 h at 37°C. Finally, luminescence was
209 integrated using a 485 ± 10 nm filter for 1 s. A SA standard curve diluted in untreated *sid2-1*
210 extract amounts ranging from 0 to 60 ng was read in parallel to allow quantification. SA
211 amounts in samples were estimated by fitting a 3rd order polynomial regression on the standards.
212
213 *Gene expression analysis*
214 Analysis of gene expression by quantitative real-time PCR was described previously (Bruessow
215 *et al.*, 2010; Gouhier-Darimont *et al.*, 2013). Briefly, tissue samples were ground in liquid
216 nitrogen, and total RNA was extracted using ReliaPrep™ RNA Tissue Miniprep (Promega)
217 according to the manufacturer's instructions. Afterwards, cDNA was synthesized from 500 ng
218 of total RNA using M-MLV reverse transcriptase (ThermoFisher) and subsequently diluted
219 eightfold with water. Reactions were performed using Brilliant III Fast SYBR-Green QPCR
220 Master Mix (Agilent) on a QuantStudio 3 real-time PCR instrument (ThermoFisher). Values
221 were normalized to the housekeeping gene *SAND* (At2g28390). Primer efficiencies were
222 evaluated by five-step dilution regression. For each experiment, three biological replicates were
223 analyzed. A list of all primers used in experiments can be found in Table S2.
224 Publically available RNA-seq data (GSE80744) from 728 accessions were recovered
225 from the Gene Expression Omnibus repository (www.ncbi.nlm.nih.gov/geo/GEO). Expression
226 data were then filtered to extract only values for accessions used in the study.
227
228 *Statistical analyses*
229 GWAS and subsequent analysis of the data obtained was performed with R software version
230 3.6. For boxplots, the thick line indicates the median, box edges represent 1st and 3rd quartile
231 respectively, whiskers cover 1.5 times the interquartile space and dots represent extreme values.
232 Boxplot width is proportional to the number of sample. When displayed, notches indicate an
233 approximate confidence interval for median values. All other analyses using mutant lines were
234 analyzed using GraphPad Prism 8.0.1.
235
236 **Results**
237 *Arabidopsis accessions display natural variation in response to P. brassicae eggs*
238 We observed that Arabidopsis accessions display varying degrees of HR-like symptoms after
239 5 days of treatment with *P. brassicae* egg extract (EE) (Fig. 1a), corresponding to the hatching
240 time of real eggs. The existence of natural variation for this defense-related trait suggests that

241 it is under genetic control, consistent with previous reports (Gouhier-Darimont *et al.,* 2013,
242 2019). While most leaves exhibited some degree of chlorosis, symptoms usually ranged from
243 no visible symptom to the formation of large patches of cell death (Fig. 1b). Moreover,
244 symptoms were most of the time restricted to the area of egg application, but could
245 exceptionally grow larger for both chlorosis and necrosis. However, because of the low
246 frequency of such events, we did not quantify those separately. This scoring scheme allows an
247 easy quantification of the HR-like (Fig. 1c). To verify that application of *P. brassicae* EE
248 mimics HR-like responses induced by real eggs, we tested whether naturally oviposited eggs
249 could also induce some degree of variation in HR-like responses. In several accessions, we
250 observed the formation of chlorotic or necrotic tissue localized around and underneath the egg
251 clutches after 5 days (Fig. 1d) and this correlated with the severity of symptoms observed after
252 EE treatment. This response was however more variable, suggesting that EE treatment might
253 be more consistent than real eggs where the protective shell may slow down the release of
254 elicitors. We thus set conditions for a GWAS using EE treatment and performed initial tests on
255 three selected accessions, using Col-0 as a control. Based on these experiments, we used short-
256 day grown plants (10L:14D), treated three leaves per plant with EE diluted 1:1 with deionized
257 water, and scored symptoms after five days (Fig. S2a-c). Because such experiment requires a
258 large amount of EE, this allowed to maximize the overall number of plant treated per accession
259 and per block, thereby improving power, even though the use of undiluted EE and/or long days
260 conditions tend to increase symptom strength (Fig. S2a-c).
261
262 *Identification of a genetic locus associated with P. brassicae egg-induced symptoms*
263 Broad sense heritability (H^2), which estimates the amount of phenotypic variance that is
264 genetically encoded, was high for symptom score (H^2 = 0.56). This finding is consistent with
265 previously reported H^2 values for defense phenotypes (Yang *et al.*, 2017). GWAS using 295
266 accessions and a fully imputed genotype matrix (Togninalli *et al.*, 2018) revealed the existence
267 of two loci associated with the degree of cell death induced by EE (Fig. 2a). One peak of
268 association was found on chromosome 2 ($-\log_{10} P = 15.96$) and constitutes the basis for another
269 study (Groux *et al.*, in preparation). Besides this peak, we observed the existence of another
270 significantly associated marker ($-\log_{10} P = 7.59$, position 16633422) on chromosome 3 (Fig.
271 2a), which explains a total of 9.56% of the phenotypic variance. By taking a closer look at this
272 genomic region, we found that this marker is located in the coding sequence of the *L-TYPE*
273 *LECTIN RECEPTOR KINASE-I.1* (At3g45330) (Fig. 2b). Interestingly, this genomic locus
274 encompasses 5 other closely related L-type clade I *LecRK* genes, *LecRK-I.2* to *LecRK-I.6*. We

275 found that high linkage disequilibrium (LD) with other surrounding markers was only observed
276 for SNPs found in the gene sequence of *LecRK-I.1* (Fig. 2b), suggesting that this gene may be
277 causal for the differences in HR-like symptoms observed. We then determined whether this
278 locus was previously identified in other publicly available GWAS. According to the araGWAS
279 database (Togninalli *et al.,* 2018), markers at the *LecRK-I.1* locus are not associated with any
280 of the defense-related or developmental phenotypes available, suggesting that it may play a
281 specific role during *P. brassicae* egg-induced response (Fig. S3).
282
283 *LecRK-I.1 is involved in EE-induced cell death*
284 In order to test whether clade I LecRKs found in this region could be involved in the response
285 to *P. brassicae* EE, we tested whether the lack of a single LecRK using T-DNA insertion lines
286 leads to alterations in different egg-induced responses. Additionally, we used CRISPR-Cas9
287 technology to delete the entire cluster of *LecRK-I.1* to *LecRK-I.6* (hereafter named *ccI.1-I.6*) in
288 order to explore whether redundancy among these genes exists (Supporting Information Fig.
289 S1). Symptom score was found to be significantly reduced in *lecrk-I.1* mutant but not in knock-
290 out lines from surrounding homologs (Fig. 3a). Consistently, deletion of the gene cluster lead
291 to symptom reduction similar to that observed in *lecrk-I*. We then quantified cell death at the
292 site of EE application by trypan blue staining and observed a similar pattern (Fig. 3b). These
293 results clearly show that *LecRK-I.1*, but not other homologs from the genomic cluster, is
294 involved in the induction of HR-like in response to EE treatment.
295 Because altered SA signaling in the *lecrk-I.1* mutant could account for the reduction in
296 cell death, we next quantified SA levels in plant treated with EE for 3 days. Total SA levels
297 were reduced in *lecrk-I.1* after 3 days of treatment compared to wild-type plants, however this
298 trend was not statistically significant (Fig. 3c). Interestingly, we observed a similar result in the
299 *lecrk-I.6* mutant and a significant reduction in *lecrk-I.3*, indicating that multiple LecRKs may
300 regulate SA accumulation. Removal of all six *LecRKs* in the *ccI.1-I.6* line resulted in
301 significantly reduced SA levels after EE treatment (Fig. 3c), further supporting that one or more
302 clade I LecRKs may participate in EE-induced SA accumulation. Altogether, these results
303 demonstrate a role for LecRK-I.1 in the regulation of cell death after insect egg perception,
304 while different clade I LecRKs regulate SA levels during this response.
305
306 *LecRK-I.1 haplotypes correlate with egg-induced symptoms*
307 The finding that natural variation in *LecRK-I.1* is associated with the inducibility of HR-like
308 responses among Arabidopsis accessions indicates that variants of this gene may alter protein

309 function(s) or gene expression. As mentioned, we observed that a single marker reached
310 genome-wide significance (SNP3; Chr. 3, 16633422) while four other SNPs had intermediate
311 *P*-values ($-\log_{10} P > 4$), all within the coding sequence of *LecRK-I.1* (Fig. 4a,b). Interestingly,
312 three out of five SNPs result in amino acid changes: one (SNP5) in the putative carbohydrate-
313 binding lectin-like domain (I228R), and two in the kinase domain (R356K, SNP3; I602V,
314 SNP1), while the two other associated markers lead to silent mutations (Fig. 4b). Additionally,
315 linkage analysis revealed that all markers are in moderate to strong LD with SNP3 (Fig. 4a).
316 Since we do not know which marker might be causal, we considered haplotypes defined by
317 SNP1-5. We next determined that at least 5 haplotypes are present in the population used for
318 GWAS mapping (Fig. 4c), with two of them being present at rather high frequency compared
319 to the other three (H2, 30.51%; H5, 58.64%; H1, H3 and H4 <7%). We observed that the
320 distribution of symptom score was not different between haplotypes having variants in the lectin
321 domain, suggesting that they might not impact this response. However, non Col-0 alleles for
322 markers in the kinase and lectin domains together are associated with significantly lower
323 symptoms (H2 vs H5, Fig. 4c). Together with the previous finding, these results potentially
324 indicate that differential kinase function may underlie natural variation in EE-induced cell
325 death. To explore this possibility, we examined whether these variants are found in functional
326 subdomains of the kinase. Based on annotation from the UNIPROT database, we found that
327 R356K lies in the predicted ATP binding site. We then performed homology-based modelling
328 using SWISSMODEL (https://swissmodel.expasy.org) to predict the 3D topology of the
329 LecRK-I.1 kinase domain in order to assess whether this amino acid change could potentially
330 alter kinase function. We found that even though R356 is indeed proximal to the ATP binding
331 site, the residue's side chain was not within 4Å of the ATP molecule and is exposed to the
332 solvent in the model obtained (P. Jimenez Sandoval, results not shown). It is thus unlikely to
333 interact with ATP within the active site. Moreover, we found that this position does not seem
334 to be conserved in the structure of immunity-related plant kinases such as BAK1, and in all
335 cases side chains were exposed to the solvent. Finally, the predicted amino acid change (R to
336 K) does not lead to significant steric or charge modification at this site. This was also true for
337 the other kinase variant, and the lectin variants appeared to be outside of the canonical lectin-
338 binding or multimerization interfaces and are also solvent exposed (P. Jimenez Sandoval,
339 results not shown). These results suggest that the identified variant may not be involved in
340 canonical kinase or receptor function based on knowledge obtained from homologous proteins.
341 We also determined whether multiple independent frameshifts and/or premature stop codon
342 may be present in natural variants at this locus, potentially leading to the production of a

343 truncated protein. We found that 18% of the sequenced accessions (23/125) possessed such
344 variants (Fig. S4a), yet this does not correlate with any difference in symptom distribution
345 independently of the haplotype considered (Fig. S4b). Moreover, premature STOP codons
346 occurred within the first 20 first amino acid of the sequence. This suggests that these frameshifts
347 and premature stops may not necessarily lead to the production of non-functional proteins,
348 possibly through the existence of alternative start sites. Collectively, our results point to a role
349 for variation in the *LecRK-I.1* gene sequence in modulating HR-like symptoms triggered by *P.*
350 *brassicae* EE. However, the link between variation in the coding of this gene and the resulting
351 phenotype is still unclear.
352
353 *LecRK-I.1 expression is haplotype-specific but does not correlate with symptom severity*
354 The absence of any obvious impact of the variation identified within the *LecRK-I.1* coding
355 sequence may indicate that an alternative process is involved. We determined whether
356 differences in gene expression may correlate with haplotype identity and HR-like symptoms by
357 measuring *LecRK-I.1* transcript levels following EE treatment in 40 accessions with high or
358 low symptoms. Gene expression was highly variable in the accessions surveyed, and expression
359 was barely detectable in certain lines (Fig. S5). We found that the two most frequent haplotypes
360 were not associated with differential *LecRK-I.1* expression after EE treatment, although one of
361 the minor haplotype displayed comparatively lower transcript levels (Fig 4). Despite this
362 haplotype-specific expression pattern, we found no correlation between symptom score and
363 *LecRK-I.1* expression levels in the surveyed accessions (Fig. 5a,b). Analysis of publicly
364 available RNA-seq data from different accessions (see methods) allowed for a similar analysis
365 using SNP3 to split the data. Consistent with our experiments, none of the alleles at this position
366 was associated with altered *LecRK-I.1* transcript levels (Fig. S6). Altogether these results
367 suggest that variation in *LecRK-I.1* expression does not account for the variation in HR-like
368 observed between accessions.
369
370 *Signatures of selection at the LecRK-I.1 locus*
371 Based on the haplotype analysis, we found that only five *LecRK-I.1* haplotypes segregate in the
372 mapping population used in this study. To get a closer look, we calculated haplotype
373 frequencies using SNP1-5 described earlier. As mentioned before, two haplotypes appear to be
374 present at high frequencies (H2, 30.51%; H5, 58.64%), while the three remaining haplotypes
375 are much less frequent (< 7%, Fig. 6a). This could indicate that haplotypes may be distributed
376 in two subpopulations due to adaptation to a type of environment for example, or that selection

377 is acting at this loci and is maintaining variation. To disentangle these hypotheses, we
378 constructed a cladogram representing genetic distance by using the kinship matrix used during
379 the GWA mapping and superimposed haplotype identity for each accession (Fig. 6b). While it
380 does appear that some low frequency haplotypes could be specific to certain geographical
381 locations, the two major haplotypes appear to have a very broad distribution in the entire
382 Arabidopsis population used for mapping (Fig. S7). The absence of any obvious link with
383 phylogenetical or geographical history suggests that selection may be acting at this locus. To
384 test the hypothesis that the *LecRK-I.1* gene sequence does not evolve neutrally, we measured
385 Tajima's D statistic along the coding sequence by using a sliding window (Fig. 6c). Tajima's
386 D compares the observed frequency of variants to those expected for a similar sequence
387 evolving neutrally (Tajima, 1989), thereby indicating any departure from neutrality. We found
388 that Tajima's D was mainly positive along the gene sequence of *LecRK-I.1*, and two short
389 stretches were found to be significantly positive (Fig. 6c). This result is indicative that *LecRK-*
390 *I.1* might not evolve neutrally. In addition, we found that nearly all genes surrounding *LecRK-*
391 *I.1* displayed negative Tajima's D values when computed over a 50 kb region (Fig. S8). Because
392 Tajima's D is sensitive to demographic history, we also performed a sliding window analysis
393 of Fu and Li's D and F statistics, which considers the species history by including an outgroup
394 (*Arabidopsis lyrata*). We observed that the pattern and sign of Fu and Li's statistics were very
395 similar to Tajima's D (Fig. 6c), further supporting the hypothesis that this locus is not evolving
396 neutrally.
397
398 *LecRK-I.1 and LecRK-I.8 function in the same pathway*
399 Based on the previously reported identification of LecRK-I.8, a close homolog of LecRK-I.1
400 (Bouwmeester & Govers, 2009; Hofberger *et al.,* 2015), as an early component of egg-triggered
401 signaling (Gouhier-Darimont *et al.,* 2019), we wondered whether both genes may function in
402 the same signaling cascade or if they are involved in parallel pathways. We crossed *lecrk-I.1*
403 with *lecrk-I.8* and measured HR-like symptoms in the single and double mutants. Both single
404 mutants displayed a similar reduction of symptoms after EE treatment. However, the double
405 mutant did not show a more severe phenotype, strongly suggesting that LecRK-I.1 and LecRK-
406 I.8 act in the same pathway (Fig. 7a). This finding is supported by the similar accumulation of
407 cell death, *PR1* transcripts and SA in single and double mutants (Fig. 7b-d).
408
409 **Discussion**

410 The recent description of LecRK-I.8 as an upstream regulator of insect egg-triggered signaling
411 opened the exciting possibility that it could be involved in egg perception (Gouhier-Darimont
412 *et al.,* 2019). However, several phenotypes including cell death were not completely abolished
413 in the *lecrk-I.8* mutant, indicating a potential redundancy with other homologs or that LecRK-
414 I.8 works as a modulator of the response. Indeed, expression all other clade I *L-type LecRKs* is
415 also induced following EE treatment (Gouhier-Darimont *et al.,* 2019). We describe here the
416 identification of LecRK-I.1 as a component of the insect egg-triggered signaling pathway in
417 Arabidopsis that specifically affects the induction of HR-like during this interaction. We show
418 that a knock-out mutant of *LecRK-I.1* leads to reduced HR-like symptoms and cell death
419 following *P. brassicae* EE treatment, while SA accumulation is reduced. Moreover, double
420 knock-out of both *LecRK-I.1* and *LecRK-I.8* did not result in a further reduction of HR-like
421 symptoms, *PR1* expression and SA accumulation, suggesting that both genes function in the
422 same transduction pathway or are part of the same complex. Other clade I *LecRK* mutants
423 showed reduced SA levels and this was further confirmed by deleting all six genes clustered in
424 the same locus. These results indicate that several clade I LecRKs are involved in the response
425 to eggs and control different aspects of this pathway. It is noteworthy that only LecRK-I.1 and
426 LecRK-I.8 affect the induction of cell death following egg perception, whereas only the *LecRK-*
427 *I.1* locus displays natural variation.
428 Despite the absence of known ligands for LecRK-I.1 and most other clade I L-type
429 LecRKs, it has been reported that LecRK-I.1 possesses a putative RGD-binding motif and that
430 it may participate in plasma membrane-cell wall interactions (Gouget *et al.,* 2006). Whether
431 insect eggs trigger changes in cell wall properties is not known, but epicuticular wax patterns
432 were shown to be altered upon *P. brassicae* oviposition on Arabidopsis (Blenn *et al.,* 2012). In
433 contrast to lectins, LecRKs show poorly conserved sugar-binding residues and it is therefore
434 unlikely that they bind carbohydrates (Bouwmeester & Govers, 2009). However, the conserved
435 hydrophobicity of the resulting cavity could recognize more hydrophobic ligands. Given the
436 lipidic nature of egg-derived inducing compounds (Bruessow *et al.,* 2010; Gouhier-Darimont
437 *et al.,* 2019), this would be consistent with a role for L-type LecRKs in egg perception.
438 Recently, the G-type LecRK LORE was found to bind to medium chain 3-OH fatty acids that
439 trigger LORE-dependent immunity (Kutschera *et al.,* 2019), further supporting the hypothesis
440 that some LecRKs may recognize hydrophobic ligands from insect eggs. Alternatively, the
441 identified LecRKs could also be involved in the perception of secondary signals such as
442 DAMPs as suggested by the identification of LecRK-I.8 and LecRK-VI.2 as potential receptors
443 for $eNAD^+$ and $eNADP^+$ (Wang *et al.,* 2017; Wang *et al.,* 2019), and LecRK-I.9 and LecRK-

444 I.5 as eATP receptors (Choi *et al.*, 2014; Pham *et al.*, 2020). Here, we provide evidence that
445 LecRK-I.1 and LecRK-I.8 function as components of the same pathway/complex. This suggests
446 that LecRK-I.1 might serve as a secondary signaling hub that controls LecRK-I.8-dependent
447 cell death. Whether LecRK-I.1 binds egg-derived elicitors alone or in association with LecRK-
448 I.8 is an intriguing hypothesis that deserves further research. A more detailed evaluation of egg-
449 triggered phenotypes in single and double mutants will help address this hypothesis.

450 We found that natural variation in the *LecRK-I.1* gene sequence was associated with
451 HR-like symptoms, consistent with the phenotype observed in the *lecrk-I.1* mutant. Two main
452 haplotypes of this gene segregate at the population level and we found that one variant, in
453 particular (SNP3), located in the kinase domain was significantly associated with HR-like
454 symptoms. Evaluation of *LecRK-I.1* expression in accessions indicates that transcript levels are
455 not associated with symptom severity, therefore suggesting that functional variation in the
456 LecRK-I.1 protein sequence modulates HR-like. Unfortunately, our analysis suggests that none
457 of the identified variants would significantly affect the lectin binding pocket or the kinase
458 domain. *In vitro* kinase assay using different variants of the kinase domain may help address
459 this question directly. Recently, LecRK-VI.2 was shown to interact with the immune co-
460 receptor BAK1 (Huang *et al.,* 2014; Wang *et al.,* 2019) and both proteins were found to be
461 necessary for $eNAD^+$- and $eNADP^+$-induced signaling. Given the critical nature of such
462 interactions for signal transduction, solvent-exposed natural variants associated with reduced
463 HR-like symptoms could affect the binding of putative co-receptors and partners of LecRK-I.1.
464 Moreover, LecRK homo- and heterodimerization was reported (Bellande *et al.,* 2017; Wang &
465 Bouwmeester, 2017), suggesting that a similar process might underlie LecRK-I.1 function and
466 such interaction may be affected in the natural variants identified.

467 At the genomic level, *L-type LecRKs* are organized in clusters of tandem repeats that
468 arose through local and whole genome duplication events (Hofberger *et al.,* 2015). Clade I *L-*
469 *type LecRK* genes appear to be mostly Brassicaceae-specific (Hofberger *et al.,* 2015; Wang &
470 Bouwmeester, 2017) and most of them have been involved in immunity. In particular, clade I
471 *L-type LecRKs* originated following the At-α whole genome duplication event (50 Mya), and
472 LecRKs that were duplicated during this period show signs of positive selection (Hofberger *et*
473 *al.,* 2015). Consistent with this study, we found that the two major *LecRK-I.1* haplotypes
474 segregate at intermediate to high frequencies, independently of any obvious geographical or
475 phylogenetical pattern. The co-occurrence of such diverged polymorphisms indicate that they
476 could be ancient and may have been maintained by selection. Consistent with the latter
477 hypothesis, we found that the *LecRK-I.1* locus does not seem to evolve neutrally, as indicated

478 by positive Tajima's D as well as Fu and Li's D/F statistics. Positive values for these statistics
479 are considered as signatures of balancing selection, a type of selection that is observed in
480 immunity-related or resistance genes for instance (Bakker *et al.,* 2006; Vila-Aiub *et al.,* 2011;
481 Huard-Chauveau *et al.,* 2013; van Velzen & Etienne, 2015; Ariga *et al.,* 2017). Balancing
482 selection collectively refers to processes by which genetic variation is maintained in a
483 population, as opposed to purifying and positive selection that reduces variation (Delph &
484 Kelly, 2014; Wu *et al.,* 2017). Different selective processes can ultimately lead to the
485 maintenance of genetic variation such as overdominance (or heterozygote advantage),
486 frequency-dependent selection or environmental heterogeneity (spatially varying selection).
487 Identification of the exact mode of balancing selection acting at this locus deserves further
488 investigation. Nevertheless, the fact that *LecRK-I.1* displays signatures of balancing selection
489 highlights the potential ecological importance of this gene in natural Arabidopsis populations.
490 While the results presented here allow to move forward in the mechanistic
491 understanding of the induction of cell death response in Arabidopsis, it is not known whether
492 this gene may also play a role in crop species such as cabbage or canola. Interestingly, induction
493 of HR-like in *B. nigra* also correlates with higher egg parasitism, indicating that both defense
494 mechanisms may be under a similar regulation (Fatouros *et al.,* 2014). Further study of this
495 region in other Brassicaceae and a screen of available germplasms for existing variation may
496 allow to demonstrate a role for clade I LecRKs in sister species.
497

498 **Acknowledgements**

499 We would like to thank Esteban Alfonso for technical help and Blaise Tissot for maintenace of
500 the plants. We thank Dany Buffat for his technical assistance during the screening. We are also
501 grateful to Nicolas Salamin and Anna Marcionetti for advices regarding the selection analyses.
502 This work was supported by the Swiss National Science Foundation (grant no 31003A_169278
503 to P.R.) and EMBO (long term fellowship ALTF 248-2018 to E.K.).
504

505 **Author Contribution**

506 RG and PR conceived and designed the research. RG and CG-D performed experiments. RG,
507 EK and PR analyzed the data. RG and PR and wrote the manuscript.
508

509 **References**

510 **Ariga H, Katori T, Tsuchimatsu T, Hirase T, Tajima Y, Parker JE, Alcázar R, Koornneef**
511 **M, Hoekenga O, Lipka AE, *et al.* 2017**. NLR locus-mediated trade-off between abiotic
512 and biotic stress adaptation in Arabidopsis. *Nature Plants* **3**: 17072.
513 **Bakker EG, Toomajian C, Kreitman, M, Bergelson J. 2006**. A genome-wide survey of R
514 gene polymorphisms in Arabidopsis. *Plant Cell* **18**: 1803–1818.
515 **Bellande K, Bono JJ, Savelli B, Jame E, Canut H. 2017**. Plant lectins and lectin receptor-
516 like kinases: How do they sense the outside? *International Journal of Molecular Sciences*
517 **18**: 1164.
518 **Bittner N, Trauer-Kizilelma U, Hilker M. 2017**. Early plant defence against insect attack:
519 involvement of reactive oxygen species in plant responses to insect egg deposition. *Planta*
520 **245**: 993–1007.
521 **Blenn B, Bandoly M, Küffner A, Otte T, Geiselhardt S, Fatouros NE, Hilker, M. 2012**.
522 Insect egg deposition induces indirect defense and epicuticular wax changes in *Arabidopsis*
523 *thaliana*. *Journal of Chemical Ecology* **38**: 882–92.
524 **Bonnet C, Lassueur S, Ponzio C, Gols R, Dicke M, Reymond P. 2017**. Combined biotic
525 stresses trigger similar transcriptomic responses but contrasting resistance against a
526 chewing herbivore in *Brassica nigra*. *BMC Plant Biology* **17**: 127.
527 **Bouwmeester K, Govers F. 2009**. Arabidopsis L-type lectin receptor kinases: Phylogeny,
528 classification, and expression profiles. *Journal of Experimental Botany* **60**: 4383–4396.
529 **Bruessow F, Gouhier-Darimont C, Bruchala A, Metraux J-P, Reymond, P. 2010**. Insect
530 eggs suppress plant defence against chewing herbivores. *Plant Journal* **62**: 876–885.
531 **Chen YH, Gols R, Benrey B. 2015**. Crop domestication and its impact on naturally selected
532 trophic interactions. *Annual Review of Entomology* **60**: 35–58.
533 **Choi J, Tanaka K, Cao Y, Qi Y, Qiu J, Lian Y, Lee SY, Stacey G. 2014**. Identification of a
534 plant receptor for extracellular ATP. *Science* **343**: 290–294.
535 **DeFraia CT, Schmelz EA, Mou Z. 2008**. A rapid biosensor-based method for quantification
536 of free and glucose-conjugated salicylic acid. *Plant Methods* **4**: 28.
537 **Delph LF, Kelly JK. 2014**. On the importance of balancing selection in plants. *New Phytologist*
538 **201**: 45–56.
539 **Desurmont GA, Weston PA. 2011**. Aggregative oviposition of a phytophagous beetle
540 overcomes egg-crushing plant defences. *Ecological Entomology* **36**: 335–343.
541 **Doss RP, Oliver JE, Proebsting WM, Potter SW, Kuy S, Clement SL, Williamson RT,**
542 **Carney JR, DeVilbiss ED. 2000**. Bruchins: Insect-derived plant regulators that stimulate

543 neoplasm formation. *Proceedings of the National Academy of Sciences USA* **97**: 6218–
544 6223.
545 **Fatouros NE, Cusumano A, Danchin EGJ, Colazza S 2016**. Prospects of herbivore egg-
546 killing plant defenses for sustainable crop protection. *Ecology and Evolution* **6**: 6906–
547 6918.
548 **Fatouros NE, Lucas-Barbosa D, Weldegergis BT, Pashalidou FG, van Loon JJA, Dicke**
549 **M, Harvey JA, Gols R, Huigens ME. 2012**. Plant volatiles induced by herbivore egg
550 deposition affect insects of different trophic levels. *PLoS One* **7**: e43607.
551 **Fatouros NE, Pineda,A, Huigens ME, Broekgaarden C, Shimwela MM, Candia IAF,**
552 **Verbaarschot P, Bukovinszky T. 2014**. Synergistic effects of direct and indirect defences
553 on herbivore egg survival in a wild crucifer. *Proceedings of the Royal Society B Biological*
554 *Sciences* **281**: 20141254.
555 **Fu YX, Li WH. 1993**. Statistical tests of neutrality of mutations. *Genetics* **133**: 693–709.
556 **Garza R, Vera J, Cardona C, Barcenas N, Singh SP. 2001**. Hypersensitive response of beans
557 to *Apion godmani* (Coleoptera: Curculionidae). *Journal of Economic Entomology* **94**: 958–
558 962.
559 **Geuss D, Stelzer S, Lortzing T, Steppuhn A. 2017**. *Solanum dulcamara*'s response to eggs
560 of an insect herbivore comprises ovicidal hydrogen peroxide production. *Plant, Cell &*
561 *Environment* **40**: 2663–2677.
562 **Gilardoni PA, Hettenhausen C, Baldwin IT, Bonaventure G. 2011**. *Nicotiana attenuata*
563 lectin receptor kinase1 suppresses the insect-mediated inhibition of induced defense
564 responses during *Manduca sexta* herbivory. *Plant Cell* **23**: 3512–3532.
565 **Gouget A, Senchou V, Govers F, Sanson A, Barre A, Rougé P, Pont-Lezica R, Canut H.**
566 **2006**. Lectin receptor kinases participate in protein-protein interactions to mediate plasma
567 membrane-cell wall adhesions in Arabidopsis. *Plant Physiology* **140**: 81–90.
568 **Gouhier-Darimont C, Schmiesing, A, Bonnet C, Lassueur S, Reymond P. 2013**. Signalling
569 of *Arabidopsis thaliana* response to *Pieris brassicae* eggs shares similarities with PAMP-
570 triggered immunity. *Journal of Experimental Botany* **64**: 665–74.
571 **Gouhier-Darimont C, Stahl E, Glauser G, Reymond P. 2019**. The Arabidopsis lectin
572 receptor kinase LecRK-I.8 is involved in insect egg perception. *Frontiers in Plant Science*
573 **10**: 623.
574 **Hilker M, Fatouros NE. 2015**. Plant responses to insect egg deposition. *Annual Review of*
575 *Entomology* **60**: 493–515.

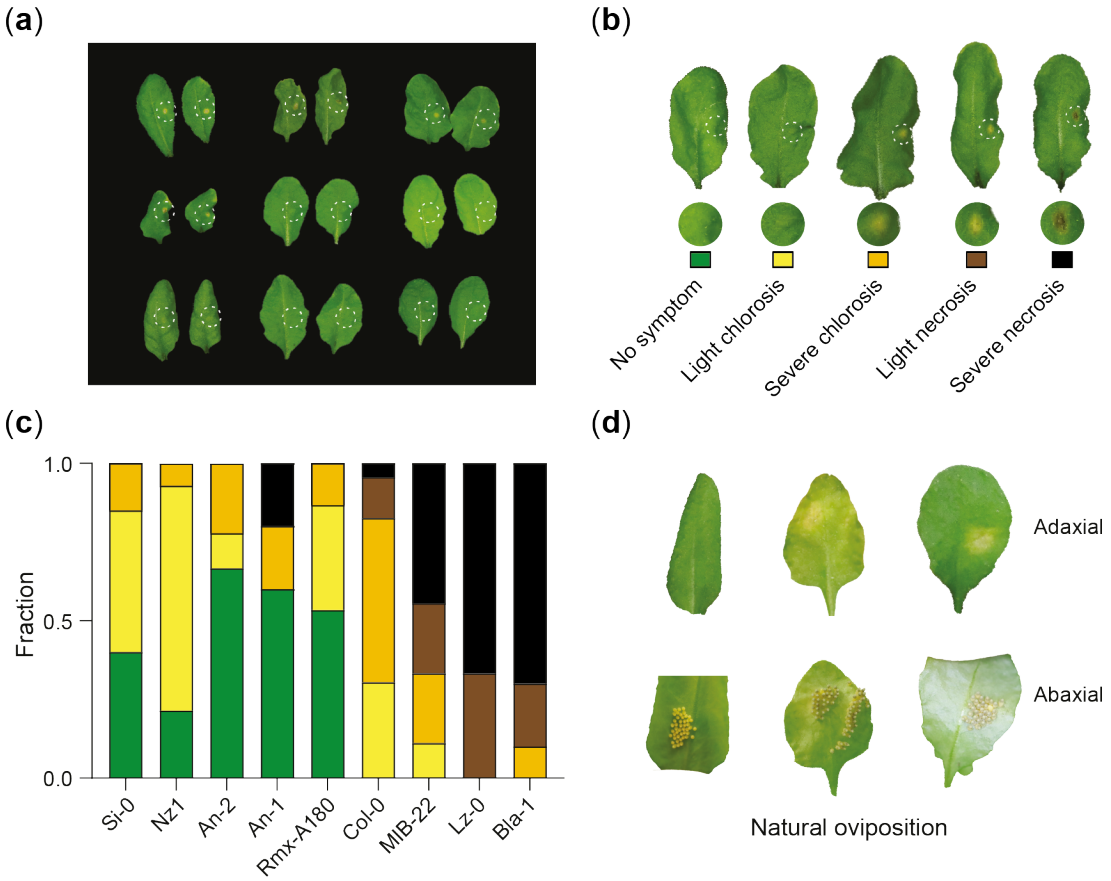
576 **Hilker M, Meiners T. 2006**. Early herbivore alert: insect eggs induce plant defense. *Journal*
577 *od Chemical Ecology* **32**: 1379–97.
578 **Hofberger JA, Nsibo DL, Govers F, Bouwmeester K, Schranz ME. 2015**. A complex
579 interplay of tandem- and whole-genome duplication drives expansion of the L-type lectin
580 receptor kinase gene family in the Brassicaceae. *Genome Biology and Evolution* **7**: 720–
581 734.
582 **Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW,**
583 **Platt A, Sperone FG, Vilhjálmsson BJ, *et al.* 2012**. Genome-wide patterns of genetic
584 variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature*
585 *Genetics* **44**: 212–216.
586 **Huang WE, Huang L, Preston M, Naylor, M, Carr JP, Li Y, Singer AC, Whiteley AS,**
587 **Wang H. 2006**. Quantitative in situ assay of salicylic acid in tobacco leaves using a
588 genetically modified biosensor strain of *Acinetobacter* sp. ADP1. *Plant Journal* **46**: 1073–
589 1083.
590 **Huang WE, Wang H, Zheng H, Huang L, Singer AC, Thompson I, Whiteley AS. 2005**.
591 Chromosomally located gene fusions constructed in *Acinetobacter* sp. ADP1 for the
592 detection of salicylate. *Environmental Microbiology* **7**: 1339–1348.
593 **Huang PY, Yeh YH, Liu AC, Cheng CP, Zimmerli L. 2014**. The Arabidopsis LecRK-VI.2
594 associates with the pattern-recognition receptor FLS2 and primes *Nicotiana benthamiana*
595 pattern-triggered immunity. *Plant Journal* **79**: 243–255.
596 **Huard-Chauveau C, Perchepied L, Debieu M, Rivas S, Kroj T, Kars I, Bergelson J, Roux**
597 **F, Roby D. 2013**. An atypical kinase under balancing selection confers broad-spectrum
598 disease resistance in Arabidopsis. *PLoS Genetics* **9**: e1003766.
599 **Kalske A, Muola A, Mutikainen P, Leimu R. 2014**. Preference for outbred host plants and
600 positive effects of inbreeding on egg survival in a specialist herbivore. *Proceedings of the*
601 *Royal Society B Biological Sciences* **281**: 20141421.
602 **Kerdaffrec E, Filiault DL, Korte A, Sasaki E, Nizhynska V, Seren Ü, Nordborg M. 2016**.
603 Multiple alleles at a single locus control seed dormancy in Swedish Arabidopsis. *eLife* **5**:
604 1–24.
605 **Kutschera A, Dawid C, Gisch N, Schmid C, Raasch L, Gerster T, Schäffer M,**
606 **Smakowska-Luzan E, Belkhadir Y, Vlot AC, *et al.* 2019**. Bacterial medium-chain 3-
607 hydroxy fatty acid metabolites trigger immunity in Arabidopsis plants. *Science* **181**: 178–
608 181.

609 **Little D, Gouhier-Darimont C, Bruessow F, Reymond P. 2007**. Oviposition by pierid
610 butterflies triggers defense responses in Arabidopsis. *Plant Physiology* **143**: 784–800.
611 **Liu Y, Wu H, Chen H, Liu Y, He J, Kang H, Sun Z, Pan G, Wang Q, Hu J, *et al.* 2015**. A
612 gene cluster encoding lectin receptor kinases confers broad-spectrum and durable insect
613 resistance in rice. *Nature Biotechnology* **33**: 301–305.
614 **Miya A, Albert P, Shinya T, Desaki Y, Ichimura K, Shirasu K, Narusaka Y, Kawakami**
615 **N, Kaku H, Shibuya N. 2007**. CERK1, a LysM receptor kinase, is essential for chitin
616 elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences USA*
617 **104**: 19613–19618.
618 **Morales J, Kadota Y, Zipfel C, Molina A, Torres M. 2016**. The Arabidopsis NADPH
619 oxidases *RbohD* and *RbohF* display differential expression patterns and contributions
620 during plant immunity. *Journal of Experimental Botany* **67**: 1663-76.
621 **Pauwels L, De Clerck R, Goossens J, Iñigo S, Williams C, Ron M, Britt A, Goossens A.**
622 **2018**. A dual sgRNA approach for functional genomics in *Arabidopsis thaliana*. *Genes*
623 *Genomes Genetics* **8**: 2603-2615.
624 **Petzold-Maxwell J, Wong S, Arellano C, Gould F. 2011**. Host plant direct defence against
625 eggs of its specialist herbivore, *Heliothis subflexa*. *Ecology and Entomology* **36**: 700–708.
626 **Pham AQ, Cho SW, Nguyen CT, Stacey G. 2020**. Arabidopsis Lectin Receptor Kinase P2K2
627 is a second plant receptor for extracellular ATP and contributes to innate immunity. *Plant*
628 *Physiology* doi: 10.1104/pp.19.01265.
629 **R Development Core Team. 2005**. R: A language and environment for statistical computing.
630 R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL:
631 http://www.R-project.org.
632 **Reymond P. 2013**. Perception, signaling and molecular basis of oviposition-mediated plant
633 responses. *Planta* **238**: 247–58.
634 **Reymond P, Weber H, Damond M, Farmer EE. 2000**. Differential gene expression in
635 response to mechanical wounding and insect feeding in Arabidopsis. *Plant Cell* **12**: 707-
636 720.
637 **Shapiro AM, DeVay JE. 1987**. Hypersensitivity reaction of *Brassica nigra* L. (Cruciferae)
638 kills eggs of *Pieris* butterflies (Lepidoptera: Pieridae). *Oecologia* **71**: 631–632.
639 **Suzuki Y, Sogawa K, Seino Y. 1996**. Ovicidal reaction of rice plants against the whitebacked
640 planthopper, *Sogatella furcifera* HORVATH (Homoptera: Delphacidae). *Applied*
641 *Entomology and Zoology* **31**: 111–118.

642 **Tajima F. 1989**. Statistical method for testing the neutral mutation hypothesis by DNA
643 polymorphism. *Genetics* **123**: 585–595.
644 **Takahashi F, Yoshida R, Ichimur K, Mizoguchi, T, Seo S, Yonezawa,M, Maruyama K,**
645 **Yamaguchi-Shinozaki K, Shinozaki K. 2007**. The mitogen-activated protein kinase
646 cascade MKK3–MPK6 is an important part of the jasmonate signal transduction pathway
647 in Arabidopsis. *Plant Cell* **19**: 805–818.
648 **Tamiru A, Khan ZR, Bruce TJ. 2015**. New directions for improving crop resistance to insects
649 by breeding for egg induced defence. *Current Opinion in Insect Science* **9**: 51-55.
650 **Togninalli M, Seren Ü, Meng D, Fitz J, Nordborg M, Weigel D, Borgwardt K, Korte A,**
651 **Grimm DG. 2018**. The AraGWAS Catalog: a curated and standardized *Arabidopsis*
652 *thaliana* GWAS catalog. *Nucleic Acids Research* **46**: D1150–D1156.
653 **Torres MA, Dangl JL, Jones JDG. 2002**. Arabidopsis gp91^{phox} homologues *AtrbohD* and
654 *AtrbohF* are required for accumulation of reactive oxygen intermediates in the plant
655 defense response. *Proceedings of the National Academy of Sciences USA* **99**: 517–522.
656 **van Velzen E, Etienne RS. 2015**. The importance of ecological costs for the evolution of plant
657 defense against herbivory. *Journal of Theoretical Biology* **372**: 89–99.
658 **Vila-Aiub MM, Neve P, Roux F. 2011**. A unified approach to the estimation and interpretation
659 of resistance costs in plants. *Heredity* **107**: 386–394.
660 **Wang C, Huang X, Li Q, Zhang Y, Li JL, Mou Z. 2019**. Extracellular pyridine nucleotides
661 trigger plant systemic immunity through a lectin receptor kinase/BAK1 complex. *Nature*
662 *Communications* **10**: 4810.
663 **Wang C, Zhou M, Zhang X, Yao J, Zhang Y, Mou Z. 2017**. A lectin receptor kinase as a
664 potential sensor for extracellular nicotinamide adenine dinucleotide in *Arabidopsis*
665 *thaliana*. *eLife* **6**: e25474.
666 **Wang Y, Bouwmeester K. 2017**. L-type lectin receptor kinases: New forces in plant immunity.
667 *PLoS Pathogens* **13**: e1006433.
668 **Wu Q, Han TS, Chen X, Chen JF, Zou YP, Li ZW, Xu YC, Guo YL. 2017**. Long-term
669 balancing selection contributes to adaptation in Arabidopsis and its relatives. *Genome*
670 *Biology* **18**: 1–15.
671 **Yamasaki M, Yoshimura A, Yasui H. 2003**. Genetic basis of ovicidal response to
672 whitebacked planthopper (*Sogatella furcifera* Horváth) in rice (*Oryza sativa* L.). *Molecular*
673 *Breeding* **12**: 133–143.
674 **Yang Y, Xu J, Leng Y, Xiong G, Hu J, Zhang G, Huang L, Wang L, Guo L, Li J, *et al.***
675 **2014.** Quantitative trait loci identification, fine mapping and gene expression profiling for

676 ovicidal response to whitebacked planthopper (*Sogatella furcifera* Horvath) in rice (*Oryza*
677 *sativa* L.). *BMC Plant Biol* **14**: 145.
678 **Yang R. 2017**. Genome-wide estimation of heritability and its functional components for
679 flowering, defense, ionomics, and developmental traits in a geographically diverse
680 population of *Arabidopsis thaliana*. *Genome* **60**: 572–580.
681 **Zvereva AS, Golyaev V, Turco S, Gubaeva EG, Rajeswaran R, Schepetilnikov MV, Srour**
682 **O, Ryabova LA, Boller T, Pooggin MM. 2016**. Viral protein suppresses oxidative burst
683 and salicylic acid-dependent autophagy and facilitates bacterial growth on virus-infected
684 plants. *New Phytologist* **211**: 1020–1034.
685

686



687

688

689 **Fig. 1** Insect egg-induced HR-like varies among natural Arabidopsis accessions. (a)
690 Representative picture of accessions (two leaves per accession) displaying varying phenotypes
691 after treatment with *P. brassicae* egg extract. (b) Representative pictures of symptoms used for
692 scoring in experiments. (c) Proportion of each symptom class developed by some accessions
693 shown in panel (a) as visualized from the adaxial side. (d) Variation in HR-like triggered after
694 natural oviposition by *P. brassicae* butterflies. In all experiments shown here, duration of
695 treatment was 5 days as this corresponds to the hatching time of naturally oviposited *P.*
696 *brassicae* eggs. The dashed circles indicate the site of treatment on the abaxial side of the leaf.
697

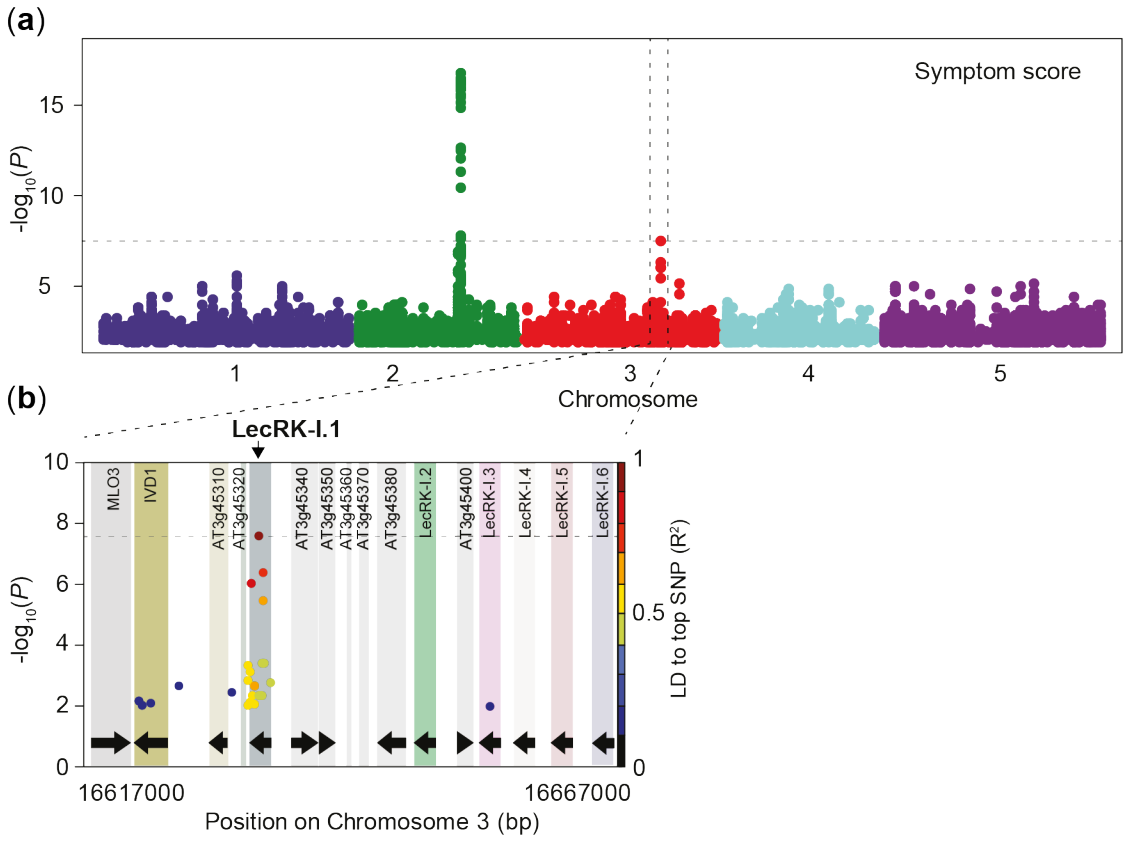
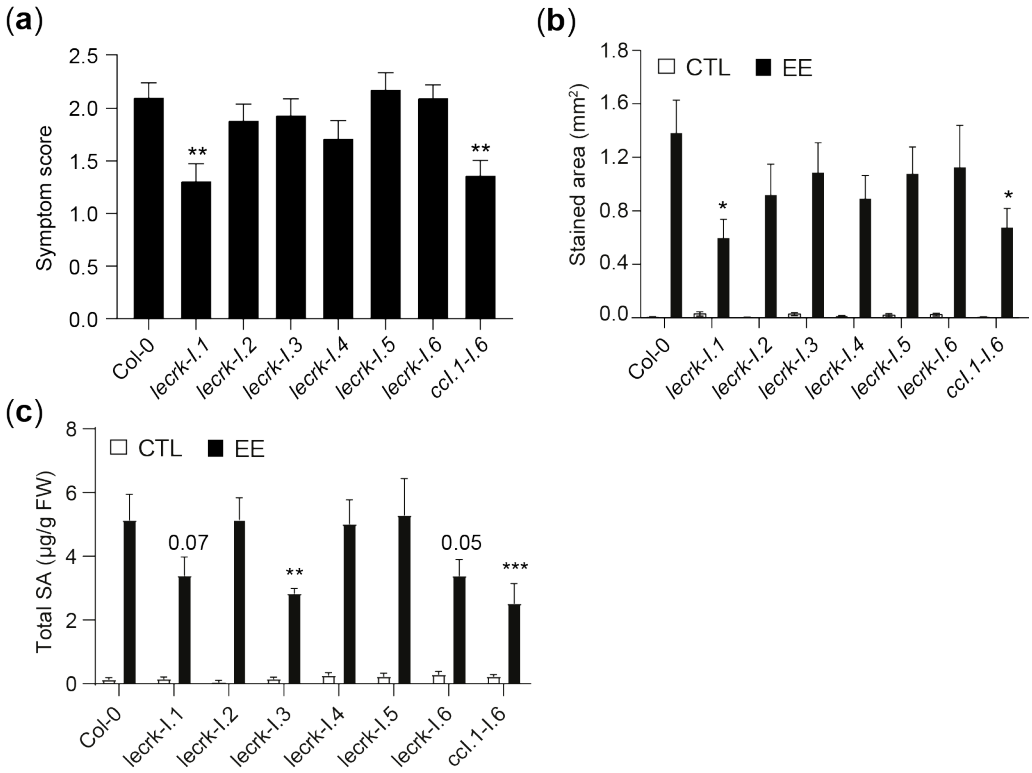


Fig. 2 Genome-wide mapping of insect-egg induced HR-like symptoms. (a) Manhattan plot of the GWAS for symptom score after 5 days of *P. brassicae* EE treatment using a linear mixed-model. Full imputed genotypes (1'769'548 SNPs) for all 295 accessions were used for mapping. Chromosomes are displayed in different colors and the dashed line indicates a significance level of 0.05 after Bonferroni correction for multiple testing. (b) Local association plot of a 50 kb region surrounding the most significant marker. The x-axis represents the genomic position on chromosome 3 and color boxes indicate genes. LD to the most significant SNP (SNP3) is indicated by a color scale.

710



711

712

713 **Fig. 3** LecRK-I.1 plays a role in the induction of HR-like symptoms following insect egg
714 perception. (a) Average symptom scores as visualized from the adaxial side after 5 days of
715 treatment with EE. Means ± SE from three independent experiments are shown (n= 12-23 for
716 each experiment). (b) Cell death as quantified by trypan blue staining after 3 days of EE
717 treatment. Untreated leaves were used as controls. Means ± SE from 8-20 leaves are shown.
718 This experiment was repeated once with similar results. (c) Total salicylic acid (SA + SAG)
719 after 3 days of EE treatment. Measurements were done on 4 plants per treatment using a
720 bacterial biosensor, untreated plants were used as control. Means ± SE of three independent
721 experiments are shown (n= 12). Stars indicate significant differences with Col-0 (ANOVA
722 followed by Dunnet's test). *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$.
723

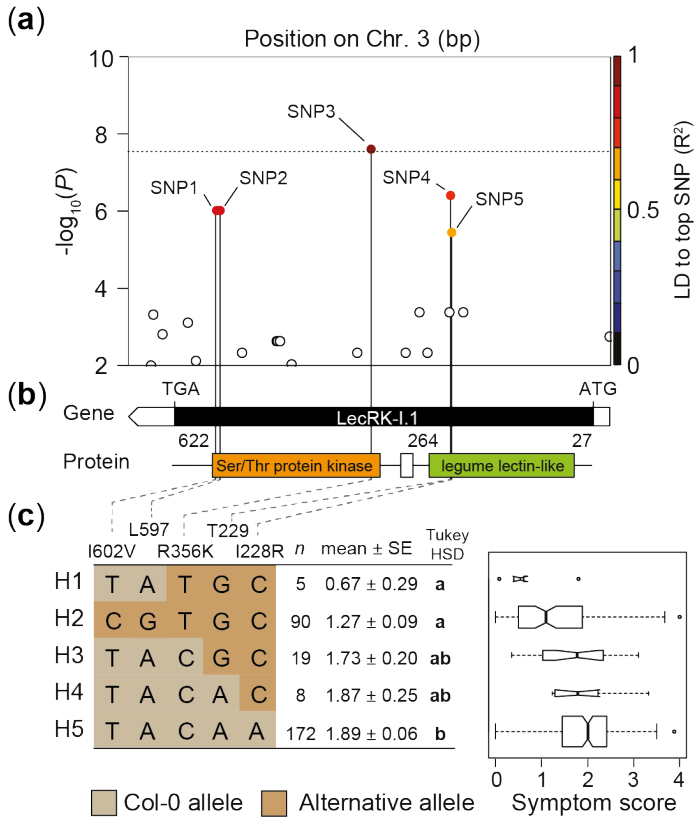


Fig. 4 Local association and haplotype analysis of the *LecRK-I.1* locus. (a) Local association plot of the *LecRK-I.1* locus. The x-axis represents the genomic position on chromosome 3. LD of markers ($-\log_{10} P>4$) to the most significant SNP (SNP3) is indicated by a color scale. The dashed line indicates the Bonferroni corrected significance threshold at $\alpha=0.05$. (b) Gene and protein domain organization according to UNIPROT. (c) Haplotype analysis using the five most significant SNPs in the *LecRK-I.1* gene. Mean symptom score ± SE is shown and n indicates the number of accessions carrying each haplotype. Different letters indicate significant difference at $P<0.05$ (ANOVA, followed by Tukey's HSD for multiple comparison). The coordinates on chromosome 3 of SNP1-5 displayed in panel (a) are the following: SNP1: 16632685; SNP2: 16632698; SNP3: 16633422; SNP4: 16633802; SNP5: 16633802.

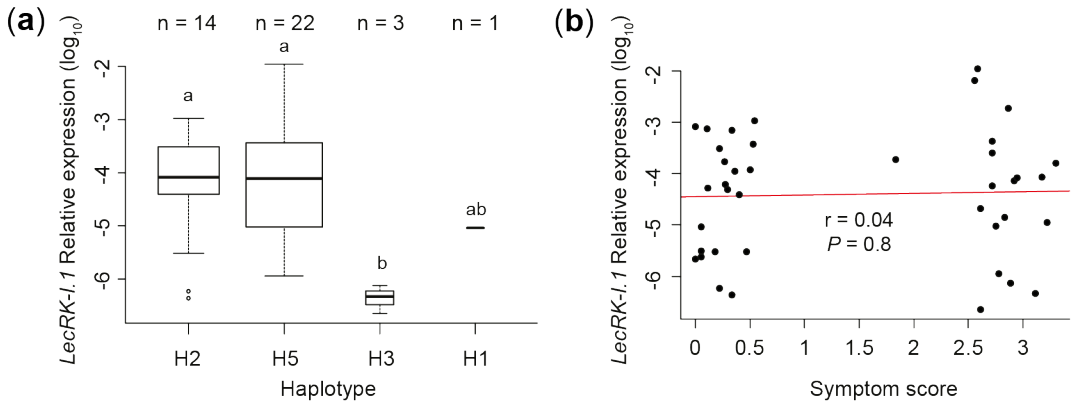
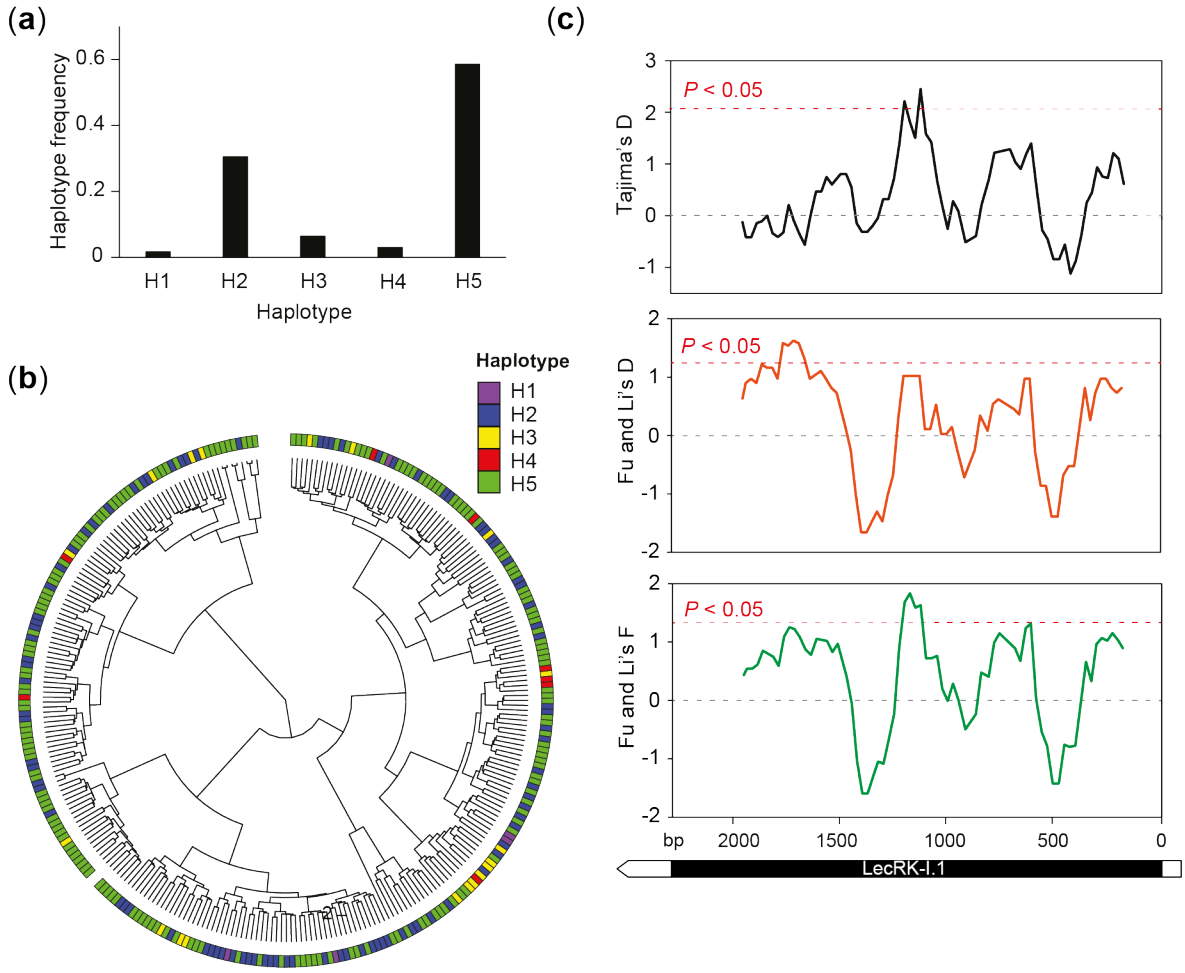


Fig. 5 Natural *LecRK-I.1* haplotypes are not differentially expressed upon EE treatment and expression does not correlate with symptom score. (a) *LecRK-I.1* expression in 40 different accessions with low and high symptoms after 72 h of EE treatment. Transcript levels were plotted according to haplotypes defined in Fig. 4. Different letters indicate significant difference at $P<0.05$ (ANOVA, followed by Tukey's HSD for multiple comparison) (b) *LecRK-I.1* expression in accessions with low or high symptoms does not correlate with symptom scores of the respective accessions. Gene expression was monitored by qPCR and target gene transcript level was normalized to the reference gene *SAND*. Means of three technical replicates are shown. Expression data were corrected by adding half the smallest non-zero value in order to avoid zero values, and $\log_{10}$-transformed prior to analysis. ns, not significant (ANOVA, $P>0.05$). Pearson correlation coefficient (r) and corresponding P value (*P*) between *LecRK-I.1* transcript level and symptom score are shown.

754



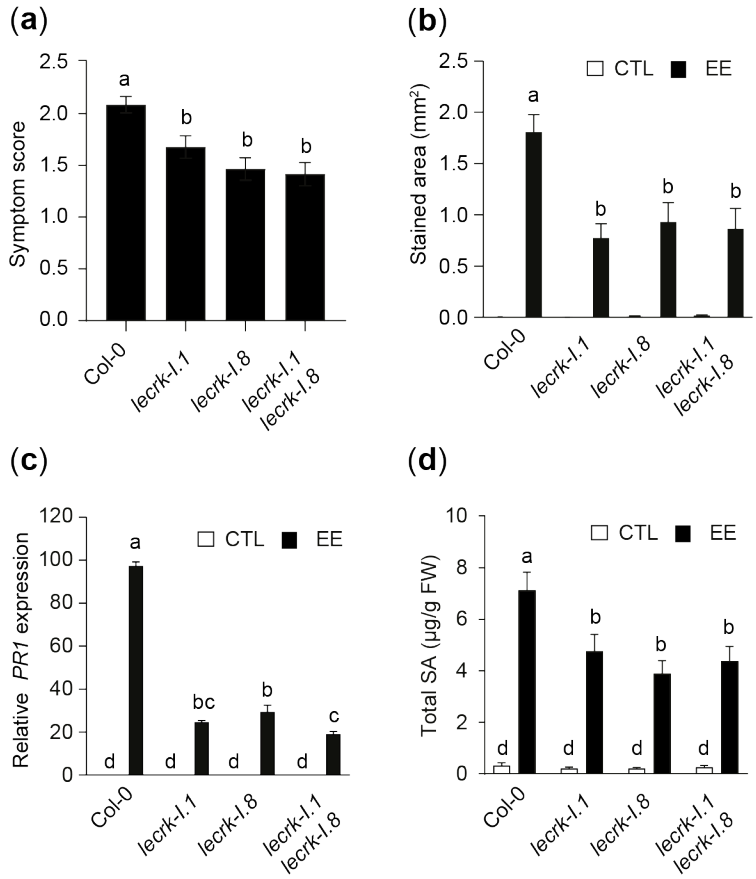
755

756

757 **Fig. 6** Signatures of selection at the *LecRK-I.1* locus. (a) Frequency of the different haplotypes
758 in the accession panel used for GWAS mapping. (b) Cladogram constructed from a genome-
759 wide kinship matrix of the 295 accessions used for GWAS. The outermost circle indicates the
760 haplotype carried by a given accession. (c) Sliding window analysis of Tajima's D, and Fu and
761 Li's D and F statistics along the *LecRK-I.1* coding region using a window size of 200 variant
762 sites and a step size of 25 sites. A subset of 125 accessions with available full genome sequences
763 was used for this analysis. The red dashed line indicates significance threshold at $P<0.05$. The
764 gene structure of *LecRK-I.1* is shown below.

765

766



767

768

769 **Fig. 7** *LecRK-I.1* genetically interacts with *LecRK-I.8*. (a) Col-0, single and double *lecrk*
770 mutants were treated with EE. Average symptom score was visualized from the adaxial side
771 after 5 days of EE treatment. Mean ± SE from three independent experiment is shown (n=12-
772 23 for each experiment). (b) Cell death as quantified by trypan blue staining after 3 days of EE
773 treatment. Untreated leaves were used as controls. Means ± SE from 8-20 leaves are shown.
774 This experiment was repeated once with similar results. (c) Expression of the marker gene *PR1*
775 after 3 days of EE treatment. Transcript levels were monitored by qPCR and normalized to the
776 reference gene *SAND*. Means ± SE of three technical replicates are shown. This experiment
777 was repeated twice with similar results (d) Total salicylic acid (SA + SAG) after 3 days of EE
778 treatment. Measurements were done on 4 plants per treatment using a bacterial biosensor,
779 untreated plants were used as control. Means ± SE of three independent experiments are shown
780 (n=12). Different letters indicate significant difference at $P < 0.05$ (ANOVA followed by
781 Tukey's HSD test for multiple comparison).