Article: Discoveries

Defects in plant immunity modulate the rates and patterns of RNA virus evolution

Rebeca Navarro[1,&], Silvia Ambrós[1,&], Anamarija Butković[1], José L. Carrasco[1], Rubén González[1], Fernando Martínez[1], Beilei Wu[1,$], and Santiago F. Elena[1,2,*]

[1]Instituto de Biología Integrativa de Sistemas (CSIC - Universitat de València), Paterna, 46182 València, Spain.

[2]The Santa Fe Institute, Santa Fe NM87501, USA.

[*]Correspondence, e-mail: santiago.elena@csic.es.

[&]Equal contribution.

[$]Permanent address: Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Haidian, 100193 Beijing, China.

Keywords: experimental evolution; generalism; plant immunity; plant-virus interactions; response to infection; specialization; virus evolution

Abstract

It is assumed that host genetic variability for susceptibility to infection necessarily conditions virus evolution. Differences in host susceptibility can either drive the virus to diversify into strains that track different defense alleles (*e.g.*, antigenic diversity) or to infect only the most susceptible genotypes. To clarify these processes and their effect on virulence, we have studied how variability in host defense responses determine the evolutionary fate of viruses. To accomplish this, we performed evolution experiments with *Turnip mosaic potyvirus* in *Arabidopsis thaliana* mutants. Mutant plants had disruptions in infection-response signaling pathways or in genes whose products are essential for potyvirus infection. Genotypes were classified into five phenogroups according to their response to infection. In order to disentangle how host susceptibility affects virus adaptation, independent viral lineages were evolved in selected plant genotypes. Evaluating disease-related traits of the evolved lineages, we found that evolution proceeded faster in the most resistant hosts than in the most permissive ones, as expected for adaptation to a harsh environment. By sequencing the genomes of the evolved viral lineages, we found that the multifunctional protein VPg turned out to be the target of selection in most host genotypes. When all evolved viral lineages were tested for fitness in all plant genotypes used in the experiments, we found compelling evidences that generalist viruses were selected by the most restrictive plant genotypes, while permissive genotypes selected for specialist viruses. Overall, this work illustrates how different host defense signaling pathways constrain not only disease-related traits but virus evolution.

Introduction

The spectrum of disease severity can be attributed to heterogeneity in virus virulence or in host factors; the two are not necessarily independent explanations and they must actually complement and/or interact with each other resulting in an arms-race coevolutionary process. A problem faced by viruses is that host populations consist of individuals that had different degrees of susceptibility to infection (Schmid-Hempel and Koella 1994; Pfenning 2001; Sallinen et al. 2020). Therefore, adaptive changes improving viral fitness in one host genotype may be selected against, or be neutral, in an alternative one. Genetic variability in susceptibility of hosts and infectiousness of viruses have been well studied in animals and plants (*e.g.*, Schmid-Hempel and Koella 1994; Altizer 2006; Hughes and Boomsma 2006; Brown and Tellier 2011; Anttila et al. 2015; Parrat et al. 2016; González et al. 2019; Sallinen et al. 2020). The interaction between host and parasite genotypes has been explained in the light of two different theoretical models that represent the two extremes on a continuum of possibilities. At the one extreme, the so-called gene-for-gene (GFG) model, in which a virus genotype exists that can infect all host genotypes and a universally susceptible host genotype should also exist (Flor 1956). Resistance occurs when a host "resistance" gene is matched by at least one virus' "avirulence" gene. Polymorphism in infectivity and resistance can be maintained only if virulence pays a cost. At the other extreme, the matching-alleles (MA) model is based on self- versus non-self-recognition systems in invertebrates. Infection is not possible unless the virus possesses all alleles that match those of the host (Frank 1993). In this case, polymorphism in infectivity and resistance are maintained by negative frequency-dependent selection.

Variability within host populations arises from nonrandom spatial distributions of genotypes: social groups of animals are more closely related to each other than to other members of the population, and plant crops are generally cultivated as genetically homogeneous plots. These situations facilitate virus transmission among genetically similar host genotypes. Spatial structure and local migration predict evolution of less aggressive exploitation in horizontally transmitted parasites (reviewed in Parrat et al. 2016). For example, Boots and Mealor (2007) observed that *Plodia interpunctella* granulosis virus evolved in spatially structured lepidopteran host populations became less virulent than the one maintained in well-mixed host populations. Berngruber et al. (2015) showed that a latent λ bacteriophage won competitions against a virulent one in a spatially structured host populations but lost in well-mixed populations. Similar results were

78 observed in laboratory evolution experiments with *Arabidopsis thaliana* and tobacco etch
79 virus (TEV): a negative association between the permissiveness to infection of different
80 plant accessions and TEV virulence was observed (Hillung et al. 2014). Finally, host
81 population structure also promotes coexistence of hosts and parasites reducing the
82 pathogens dispersal (Brockhurst et al. 2006).

83 Theoretically, in the absence of host heterogeneity, parasites must evolve toward a
84 host exploitation strategy that maximizes transmission with low virulence (Haraguchi and
85 Sasaki 2000; Regoes et al. 2000; Rodríguez and Torres-Sorando 2001; Ganusov et al.
86 2002; Gandon 2004; Lively 2010; Moreno-Gámez et al. 2013). Chabas et al. (2018)
87 experimentally showed that evolutionary emergence is more likely when the host
88 population contains intermediate levels of resistant hosts, confirming this prediction using
89 a set of phages and bacterial hosts with alterations in the CRISPR/Cas-mediated
90 immunity.

91 Viral infection of plants is a complex system in which the virus parasitizes the host
92 and utilizes all its cellular resources to replicate and systemically spread. In response,
93 plants have evolved intricated signaling mechanisms that limit the spread of the virus,
94 resulting in resistance (Zhou and Zhang 2020). A variety of factors contribute to plant
95 resistance to viral infections. Broadly speaking, these factors can be classified into *basal*
96 if they are pre-existing and limit within-cell propagation and cell-to-cell spread, and
97 *inducible* if they are only activated upon infection and inhibit systemic virus movement
98 and replication. Basal mechanisms include susceptibility (*S*) genes that involve alleles of
99 cellular proteins that do not interact properly with viral factors; *e.g.*, translation initiation
100 factors required by the virus for successful exploitation of the cell's protein synthesis
101 machinery, heat shock proteins that assist the formation of multiprotein complexes, or
102 DNA binding phosphatases (Carr et al. 2010; Mäkinen 2019). In contrast, inducible
103 mechanisms include genes whose expression results in a broad-scale change in plant
104 physiology via diverse signal transduction pathways, particularly those regulated by the
105 hormones salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) (Soosaar et al. 2005;
106 Carr et al. 2010). These changes include local cell apoptosis (*e.g.*, hypersensitive
107 responses - HR; Loebenstein 2009), the upregulation of nonspecific responses against
108 many different types of pathogens throughout the entire plant (systemic acquired
109 resistance - SAR- and induced systemic resistance - ISR) (Kachroo et al. 2006; Carr et al.
110 2010), and the activation of the RNA-silencing-based resistance, which seems to play a
111 role both in basal and inducible mechanisms (Voinnet 2001; Carr et al. 2010). Early host

112 responses following virus detection include changes in ion fluxes (mainly Ca^{++}),
113 activation of signaling pathways, major alterations of transcriptomic profiles, generation
114 of reactive oxygen species (ROS) and production of nitric oxide (NO) (Soosaar et al.
115 2005). These immediate changes are followed by HR and the recruitment of SA and
116 JA/ET signaling pathways. The SA-mediated defense signaling pathway results in SAR,
117 while the JA/ET-mediated defense signaling pathway results in ISR, the latter being
118 specifically involved in interactions between plants and beneficial microbes. Indeed, it
119 appears that ISR is not effective against infection by most viruses (Ton et al. 2002;
120 Loebenstein 2009; Pieterse et al. 2009). Both SAR and ISR pathways converge into two
121 master regulators, the *ENHANCED DISEASE SUSCEPTIBILITY 1* (*EDS1*) and the
122 *PHYTOALEXIN DEFICIENT 4* (*PAD4*) genes; *EDS1* and *PAD4* repress ISR and promote
123 SAR. Although the SA and JA/ET pathways have been classically viewed as mutually
124 antagonistic, several studies have revealed positive and negative crosstalk between them
125 (van Wees et al. 2000; Pieterse et al. 2012) as well as with the RNA-silencing pathway
126 (Soosaar et al. 2005; Carr et al. 2010; Yang et al. 2020). This crosstalk, which has the
127 network topological structure of an incoherent feed-forward loop, creates robustness and
128 tunability in the plant immune network (Mine et al. 2017).

129 An open question is how the intricacies of the underlying defense regulatory
130 pathways of *A. thaliana*, the existence of multiple defense responses and their potential
131 crosstalk determines the evolutionary fate of viruses. For instance, would two
132 independent defense pathways select for specialist viruses adapted to counteract each one
133 thus conforming a fitness tradeoff? Would a tradeoff in the investment of plant resources
134 between pathways (*e.g.*, SAR *vs* ISR) result in a diversification of virus adaptive
135 strategies? Would viruses targeting crosstalk points across pathways evolve as
136 generalists? Here we will use turnip mosaic virus (TuMV; species *Turnip mosaic*
137 *potyvirus*, genus *Potyvirus*, family *Potyviridae*) - *A. thaliana* experimental pathosystem
138 to explore these questions. After exploring the variability in phenotypic responses to
139 TuMV infection of a collection of *A. thaliana* mutants in basal and inducible resistance
140 mechanisms, we choose nine mutants which cover the entire spectrum in phenotypic
141 responses. Then, we performed evolution experiments on each plant mutant genotype
142 and track the evolution of several disease-related traits. At the end of the evolution
143 experiment, we evaluated the effect of host genotypes in the rates of virus evolution and
144 the contribution of historical contingency, selection and stochasticity into the outcome of
145 evolution. Next, we explored whether the different TuMV lineages evolved as specialist

146 or generalists depending on the disrupted defense mechanism of their local hosts. Finally,
147 we sought to identify the molecular changes experienced in the genome of the different
148 viral lineages and to explore the possible adaptive value of a few convergent mutations.
149
150

151 # Results

152 **Classification of *A. thaliana* genotypes according to their phenotypic response to**
153 **TuMV infection**

154 The 22 *A. thaliana* mutant genotypes used in this study are shown in table 1, including
155 information about the affected signaling pathways or cellular processes as well as the
156 expected phenotype of infection relative to wildtype (WT) plants based on the description
157 of these mutants. Genotypes were classified according to their phenotypic response to
158 TuMV infection based in five different disease-related traits measured 18 days post-
159 inoculation (dpi). The disease-related traits measured were change in dry weight (Δ*DW*),
160 disease progression (area under the disease progress step curve or *AUDPS*), infectivity
161 (*I*), symptoms severity (*SS*), and viral load (*VL*). Figure 1 shows the nearest-neighbor
162 clustering of genotypes according to their multi-trait phenotypic response to TuMV
163 infection (supplementary figure S1, Supplementary Material online); we found five
164 significant phenogroups (hereafter named as G1 to G5). For all four members of G1 the
165 response was consistent with an enhanced resistance response to TuMV infection: no
166 significant Δ*DW* whilst a significant and consistent reduction in *AUDPS*, *I*, *SS*, and *VL*
167 relative to infected WT plants. Three members of G1 had enhanced SAR response (table
168 1) and the fourth one (*i4g2*) is a well-known *S* gene involved in plant resistance to
169 potyviruses (Nicaise et al. 2007; Gallois et al. 2010). The only member of G2 was the
170 strong apoptosis-inducer $p58^{IPK}$, which shows no significant changes in *AUDPS*, *I*, *SS*,
171 and *VL* relative to infected WT plants but a significant reduction in Δ*DW*. Infected $p58^{IPK}$
172 plants were heavier than the noninfected controls which points to a case of increased
173 tolerance to infection. The only member of G3 was *dbp2*, that specifically increased
174 resistance against another potyvirus (Castelló et al. 2011). This mutant showed a very
175 interesting response to TuMV infection: Δ*DW* was significantly increased while *AUDPS*,
176 *I* and *SS* were significantly reduced; no effect in *VL* was observed. Thus, *dbp2* represents
177 an intermediate response to TuMV infection between the highly resistant G1 and those
178 genotypes in G4. G4 represents a sort of hotchpotch formed by genotypes that do not

179 show a clear phenotypic difference from WT. Remarkably, though, with the exception
180 of *cpr5-2*, *dcl2 dcl4*, and *dip2*, all show significant reductions in *VL*. Finally, *jin1* was
181 the only member of G5, which shows a significant increase in all measured traits except
182 *VL*, consistent with and enhanced susceptibility to TuMV infection. This is a surprising
183 finding since *jin1* has been described as inducing a constitutive expression of the SA-
184 dependent defenses that made plants extremely resistant to the infection of biotrophic
185 bacterial pathogens (Laurie-Berry et al. 2006). To further characterize the progression of
186 infection, we selected a subset of representatives from groups G1 (*eds8-1*) and G4 (*cpr5-*
187 *2*, *dcl2 dcl4*, *hsp90-1*, and WT) and evaluated daily *I*, *SS* and *VL* (supplementary figure
188 S2, Supplementary Material online). The *cpr5-2* plants started showing symptoms faster,
189 though only 80% of plants showed symptoms after 15 dpi. Only 20% of *eds8-1* plants
190 showed symptoms of infection, which were delayed. Infected *hsp90-1* and *dcl2 dcl4*
191 plants were undistinguishable from the infected WT ones. Regarding *VL*, all genotypes
192 accumulated less virus than WT plants, amongst them *cpr5-2* accumulated the least
193 amount of virus (supplementary figure S2, Supplementary Material online).
194 Representatives from the five groups were selected for the subsequent evolution
195 experiment (shadowed rows in table 1): *eds8-1* and *i4g2* from G1 representing the hardest
196 selection regimes (more resistant genotypes); $p58^{IPK}$ from G2; *dbp2* from G3; *cpr5-2*,
197 *dcl2 dcl4*, *hsp90-1*, *npr1-1*, and WT from G4; and *jin1* from G5 representing the softest
198 selective regime (less resistant genotype).
199
200 **Experimental evolution in representative *A. thaliana* genotypes**
201 First, we sought to evaluate the dynamics of evolution for *AUDPS* (figure 2), *I* (figure 3)
202 and *VL* (figure 4). Data in these figures were fitted all together into the multivariate
203 analysis of covariance (MANCOVA) model described in equation 1 of the Materials and
204 Methods section. Notice that lineage *cpr5-2*/L2 showed a quite different evolution
205 pattern from the other four lineages evolved in *cpr5-2*. It never increased in either of the
206 three traits and was lost after passage five (figures 2 and 3). Therefore, we ended up with
207 46 evolved lineages (two in the WT, four in *cpr5-2* and five in each of the other eight
208 plant genotypes).
209 Table 2 shows the results of this MANCOVA. Despite a considerable amount of
210 noise in the time series, we found some significant results that can be summarized as
211 follows. First, a net significant and large (η_P^2 = 0.735) effect associated to the
212 evolutionary passages (*t*) has been observed, indicating that the three phenotypic traits

213 related to TuMV infection evolved during the experiment. Second, plant genotypes (G)
214 had a highly significant and large ($\eta_P^2 = 0.547$) effect on the phenotypic traits, suggesting
215 that TuMV evolutionary dynamics were strongly influenced by the local host genotype
216 in which it was being passaged. More interestingly, this effect was strongly dependent
217 on the passage number ($t \times G$; $\eta_P^2 = 0.463$), thus suggesting that the slope of the relationship
218 between phenotypic traits and time was influenced by the local host genotype. This effect
219 will be further evaluated in the next section. Third, independent lineages evolved in the
220 same host genotype show a strong degree of evolutionary parallelism, as indicated by
221 nonsignificant differences among viral lineages within local host genotype (factors $L(G)$
222 and $t \times L(G)$). However, this last conclusion should be taken with certain caution since the
223 power of the two tests is relatively low ($1 - \beta < 0.800$) and hence we might be wrongly
224 failing to reject the null hypothesis (type II error).

225 A relevant question in evolutionary biology is the extent in which ancestral
226 differences determine the fate of evolution. In other words, the relative contribution of
227 adaptation, chance and historical contingency to the evolution or organismal fitness.
228 Following the logic exposed by Travisano et al. (1995), we can imagine two situations.
229 First, ancestral statistical differences among phenotypes could be preserved along
230 evolution despite a net increase in the mean trait values (due to selection) and differences
231 among replicated lineages (due to chance; *i.e.*, mutation and drift). In this situation, we
232 should expect a nonzero slope in a regression of the evolved phenotypic values against
233 the ancestral ones. The closer the slope to a value of one, the more importance of ancestral
234 differences. Second, if initial trait variation among ancestral genotypes was eliminated
235 from the evolved populations because the combined effect of adaptation and chance, we
236 should expect a regression slope smaller than one. The less effect of ancestral differences,
237 the flatter the slope, being zero in the extreme case where ancestral differences are
238 completely erased. To evaluate the role of the ancestral differences in *AUDPS*, *I* and *VL*
239 of TuMV across the nine different plant genotypes, we evaluated the magnitude of their
240 change at the end of the evolution experiment. Figure 5 shows the plots of evolved *vs*
241 ancestral values. For illustrative purposes, the solid diagonal lines represent the null
242 hypothesis of absolute preservation of ancestral differences. In the case of *AUDPS* (figure
243 5A), a significant regression exists between evolved and ancestral values ($R = 0.491$, $F_{1,44}$
244 $= 14.000$, $P < 0.001$), with slope 0.284 ±0.076 (±1 SD). Since the slope is still
245 significantly different from zero, yet clearly flatter than the diagonal ($t_{44} = 9.422$, $P <$

246 0.001), we conclude that ancestral differences have been mostly removed by the
247 combined action of selection and chance, yet not completely erased. In the case of *I*
248 (figure 5B), the regression was not significant ($R = 0.094$, $F_{1,45} = 0.405$, $P = 0.528$), thus
249 suggesting that for this trait ancestral differences among *A. thaliana* genotypes have been
250 fully erased by adaptation and chance. Finally, in the case of *VL* (figure 5C; WT not
251 included in this analysis), the regression of evolved values on the ancestral ones was
252 significant ($R = 0.417$, $F_{1,43} = 8.822$, $P = 0.005$), suggesting again that ancestral
253 phenotypic differences have been largely removed by selection and chance yet some still
254 persist (regression slope -0.142 ± 0.048 significantly different from one: $t_{42} = 23.903$, P
255 < 0.001).
256
257 **Rates of phenotypic evolution for disease-related traits**
258 As mentioned in the previous section, we have observed a significant effect of *A. thaliana*
259 genotypes in the temporal evolution of the three phenotypic traits. To further explore this
260 effect, we estimated the rates of phenotypic evolution of *AUDPS* and *I* using the
261 ARIMA(1,0,0) model described by equation 2 in Materials and Methods. *VL* was not
262 included in this analysis since the number of data points (three) in the time series was the
263 same as the number of parameters to be estimated, thus likely incurring in overfitting
264 problems.
265 Figure 6 shows the estimated rates of evolution (ν) for host genotypes ordered from
266 the most resistant (G1) to the most permissive (G5) ones. Very interestingly, a significant
267 negative regression coefficient exists for both traits [*AUDPS*: $-1.313 \cdot 10^{-2} \pm 0.383 \cdot 10^{-2}$
268 ($R = 0.459$, $F_{1,44} = 11.724$, $P = 0.001$) and *I*: $-2.576 \cdot 10^{-3} \pm 0.790 \cdot 10^{-3}$ ($R = 0.4413$, $F_{1,44}$
269 $= 10.621$, $P = 0.002$)], suggesting that evolution always proceeded faster in the most
270 restrictive hosts and slower in the most permissive ones.
271
272 **Dissecting the role of specific defenses in TuMV evolution**
273 The rates of TuMV phenotypic evolution estimated in the previous section were further
274 analyzed in the context of the mutated defense signaling pathways or recessive
275 resistances. Rates of evolution ν_{AUDPS} and ν_I were fitted to the MANOVA model shown
276 in equation 3 of the Materials and Methods section. Figure 7 graphically summarizes the
277 results of these analyses. First, we sought for differences in rates of evolution between
278 permissive (soft selection) and restrictive (hard selection) hosts (figure 7A). The

279 multivariate analysis showed that lineages evolved in the more restrictive hosts evolved
280 significantly faster than those evolved in the more permissive hosts (Λ = 0.780, $F_{2,43}$ =
281 6.080, P = 0.005). This effect was mostly driven by differences in the rate of infectivity
282 evolution v_I, as shown by the corresponding univariate analyses, being on average 3.14
283 times faster in the most restrictive hosts. This observation confirms the results shown in
284 the previous section.

285 Next, we evaluated whether differences exist between the rates of phenotypic
286 evolution for viruses evolved in host genotypes carrying mutations affecting basal or
287 inducible resistances (figure 7B; WT data for this analysis excluded). In this case, the
288 multivariate analysis found a small, yet significant effect (Λ = 0.790, $F_{2,34}$ = 4.507, P =
289 0.018), though none of the two univariate analyses had enough power to detect it (indeed,
290 changes went in opposite directions: 1.02-fold larger for v_{AUDPS} in the plants with
291 mutations in basal resistances but 1.31 times larger for v_I in the plants with mutations in
292 the inducible resistances).

293 Figure 7C shows the comparation of rates of TuMV phenotypic evolution for
294 lineages evolved in plant genotype mutants for the SA-dependent signaling defense
295 pathways. In this case we have three categories compared to WT: mutations inducing a
296 down-expression of the pathway, mutations with no effect on this pathway and mutations
297 resulting in overexpression. Again, the MANOVA found highly significant differences
298 among the three categories (Λ = 0490, $F_{4,84}$ = 8.996, P < 0.001). The univariate analyses
299 showed that most of the differences were due to a 2.38 times faster evolution of I in the
300 host genotypes over-expressing the SAR defenses, than in those with a WT-like or even
301 down-expression of the SA signaling which results in a reduced SAR.

302 SA- and JA/ET-dependent pathways have been considered as antagonistic defenses
303 (see Introduction), although more recent evidence points toward an active crosstalk
304 between both (van Wees et al. 2000; Pieterse et al. 2012; Mine et al. 2017). Under the
305 antagonistic hypothesis, we should expect the opposite trend among lineages evolved in
306 plants with affected JA/ET signaling than the one observed for the lineages evolved in
307 plants with affected SA signaling (figure 7C). Figure 7D shows the data of evolution
308 rates classified according to whether host genotypes had mutations affecting the JA/ET
309 defense signaling pathway that resulted ISR. The MANOVA found a significant
310 difference between the two categories (Λ = 0.766, $F_{2,43}$ = 6.580, P = 0.003), entirely
311 driven by differences in v_I, as confirmed by the univariate analyses. Indeed, as expected,

312 the rates of evolution for infectivity were 2-fold faster for TuMV lineages evolved in the
313 ISR deficient hosts, in agreement with the SAR/ISR antagonistic hypothesis.
314 Finally, we tested the effect of knocking down the host's RNA-silencing pathway
315 in the rates of TuMV evolution (figure 7E). In this case, the multivariate analysis also
316 found highly significant differences between rates of viral evolution in plants with a fully
317 functional RNA-silencing pathway and the silencing-deficient *dcl2 dcl4* mutant (Λ =
318 0.850, $F_{2,43}$ = 3.801, P = 0.030). The average v_I values were 21-fold larger in plants
319 expressing a fully functional RNA-silencing pathway than in *dcl2 dcl4* plants unable of
320 siRNA production.
321 In conclusion from this section, we found that regardless of the defense signaling
322 pathway, viruses evolved in permissive plants always evolve slower than their
323 counterparts evolved in more restrictive plant genotypes.
324
325 **Permissiveness of *A. thaliana* genotypes to infection drives the evolution of TuMV**
326 **lineages into specialists or generalists**
327 To analyze the specificity of adaptation of each evolved TuMV lineage, we performed a
328 full cross-infection experiment. In this experiment, the 44 lineages evolved in the mutant
329 genotypes (excluding the two WT-evolved lineages) were inoculated into ten plants of
330 the nine *A. thaliana* mutant genotypes used in the evolution experiments (excluding WT
331 plants). The presence of symptoms in each plant was recorded daily for up to 12 dpi.
332 This dataset was analyzed using two different approaches, first an ANOVA-like that
333 evaluates the effect of host genotypes, viral lineages and their interactions (Schmid-
334 Hempel 2011); and second, the inference and characterization of nestedness, modularity
335 and specialization of a bipartite infection network (Weitz et al. 2013). For the first
336 approach we fitted the presence/absence of symptoms data to the logistic regression
337 model shown in equation 4 of the Materials and Methods section. The results are shown
338 in table 3. A test host genotype (*TG*) factor describes the effect of differences among host
339 genotypes, irrespective of the infecting viral lineage; and thus, characterizes whether
340 some host genotypes are more susceptible than others to infection. As shown in table 3,
341 *TG* has a highly significant effect, both by itself and in the interaction with *t*. The
342 magnitude of the net effect was very large (η_P^2 = 0.569), while its interaction with *t* was
343 rather small in magnitude (η_P^2 = 0.010). On average, the most susceptible host genotypes
344 were *jin1*, *dcl2 dcl4* and *cpr5-2* ($\bar{f}$ = 1.000), whereas the most resistant host genotypes

345 were *p58*^{IPK} ($\bar{f}$ = 0.459 ±0.013) and *eds8-1* ($\bar{f}$= 0.233 ±0.013). Similarly, a highly
346 significant local host genotype (*LG*) effect exists by itself (of large magnitude η_P^2 = 0.288)
347 as well as in combination with *t* (though the magnitude of the interaction was rather small
348 in magnitude η_P^2 = 0.005), which means that the observation depends on the viral lineage
349 independently of the host genotype. On average, the lineages evolved in *eds8-1* ($\bar{f}$ =
350 0.997 ±23.710) and *dbp2* ($\bar{f}$ = 0.994 ±26.108) induced symptoms faster; while in contrast,
351 lineages evolved in *i4g2* ($\bar{f}$ = 0.788 ±279.090) showed symptoms much slower. Finally,
352 and most interestingly, a highly significant interaction term *TG*×*LG* (as well as in
353 combination with *t*) indicates that the outcome depends on the particular combinations of
354 host genotypes and TuMV lineages. Indeed, the magnitude of the interaction effect was
355 large (η_P^2 = 0.314), though its dependence with dpi was not so relevant (η_P^2 = 0.048). For
356 example, in the most susceptible genotypes, where *cpr5-2* perfectly illustrates this case:
357 lineages evolved in *cpr5-2*, *dbp2*, *dcl2 dcl4*, *eds8-1*, and *p58*^{IPK} all show $\bar{f}$ = 1.000,
358 whereas lineages evolved in *i4g*2 and *jin1* show much lower $\bar{f}$ values (0.674 ±0.028 and
359 0.849 ±0.032, respectively). In the opposite extreme, *e.g.*, *eds8-1*-evolved lineages show
360 $\bar{f}$ = 0.575 ±0.050 while *jin1*-evolved ones show $\bar{f}$ = 0.124 ±0.030; while all other lineages
361 show intermediate $\bar{f}$ values.

362 For the second analytical approach, we used the full cross-infection dataset to
363 estimate the corresponding *AUDPS* values, thus creating an infection matrix of 44 rows
364 [*L*(*LG*)] by nine columns (*TG*). This matrix was transformed into a new binary matrix in
365 which a value of 1 meant that the *AUDPS* for a given TuMV lineage, in a particular *TG,*
366 was greater or equal than the one observed for its *LG* and 0 otherwise. Figure 8A shows
367 the packed infection matrix with rows and columns organized to better highlight its
368 nestedness and modularity (Weitz et al. 2013). The matrix could be also represented as
369 a bipartite infection network (figure 8B) that provides the same information but in a more
370 graphical way. The infection network was significantly nested (T = 14.706, P < 0.001),
371 with the most permissive genotypes *jin1* and *dcl2 dcl4* being susceptible to most viral
372 lineages (universally susceptible hosts), while the most resistant genotype *eds8-1* and
373 *p58*^{IPK} were only successfully infected by some lineages that were precisely evolved in
374 these two genotypes. Likewise, the most generalist viral lineages were *eds8-1*/L4,
375 *p58*^{IPK}/L3 and *p58*^{IPK}/L4 that infected the nine mutant host genotypes (figure 8) with
376 equal efficiency (universally virulent viruses). On another hand, the most specialized
377 viral lineages that were able to successfully infect only their local hosts, were all surviving

lineages evolved in *cpr5-2*, four lineages evolved in *dcl2 dcl4* and four lineages evolved in *jin1* (figure 8). Therefore, we conclude that local host's resistance to infection positively correlates with the host range of the evolved lineages: more resistant host genotypes selected for very generalist viruses while less resistant host genotypes selected for very specialized viruses, in agreement with a GFG infection model (Weitz et al. 2013).

Next, we computed Blüthgen et al. (2006) specialization indexes for the packed matrix (figure 8). Firstly, we evaluated the standardized species-level measure of partner diversity *d'*. *d'* ranged from zero for the most generalist TuMV evolved lineages (*eds8-1*/L4, $p58^{IPK}$/L3 and $p58^{IPK}$/L4) to one for the most specialist ones (the four surviving *cpr5-2*-evolved lineages, *dbp2*/L2, *dcl2 dcl4*/L2 - L5 and *jin1*/L1 - L4) (last column in figure 8A). Secondly, we computed the network-level specialization index H_2', obtaining a value of $H_2' = 0$ (uncorrected $H_2 = 5.024$, $H_{min}' = 3.534$, $H_{max}' = 5.024$; $P < 0.001$), which means that, overall, the binary infection network shows a degree of specialization. Most of the lineages were able of infect one or a few host genotypes, a result mainly driven by those lineages evolved in the less resistant host genotypes.

Finally, we sought to explore the modularity of the infection network. Indeed, from an ecological standpoint, a significant network-level degree of specialization means that some TuMV lineages do not interact well with some of the *A. thaliana* genotypes, thus creating the possibility of modularity. In this context, a module will refer to an aggregated set of viral lineages and their hosts characterized by more interactions within the module than between modules (Newman 2006; Dormann and Strauss 2014). To this end, we computed the *Q* modularity index, which ranges from zero, when the community has no more links within modules than expected by chance, to a maximum value of one. We found a small yet significant modularity ($Q = 0.234$, $P = 0.047$) in the infection matrix (figure 8A). What are the implications of modularity in TuMV pathogenesis? Modularity is expected for a MA infection model (Weitz et al. 2013). More similar host genotypes, *e.g.*, those with mutations affecting the same signaling pathway, may be selecting for viral lineages with similar properties, thus being able of infect a subset of plant genotypes with equal efficiency. An alternative hypothesis that can explain modularity may be parallel evolution of lineages evolved into the same host genotype; the most representative cases of this possibility are the lineages evolved in *cpr5-2*, *dcl2 dcl4* and *jin1*. To distinguish between these two hypotheses, we computed a reduced matrix averaging the observed *AUDPS* among lineages evolved in the same host genotype and transforming this nine-by-nine infection matrix into its corresponding packed binary matrix, as described above.

412 If the reduced matrix still shows significant modularity, the first hypothesis will hold on,
413 but if modularity vanishes by averaging lineages evolved in the same host genotype, then
414 the alternative hypothesis will be more parsimonious. The reduced matrix is still
415 significantly nested ($T = 13.064$, $P < 0.001$), but modularity was not significant anymore
416 ($Q = 0.286$, $P = 0.137$). This supports the hypothesis that the observed modularity was
417 driven by convergent phenotypic evolution of lineages evolved into the same host
418 genotype rather than by the overlap of selective pressures by genotypes carrying
419 mutations in the same defense mechanism.
420

421 **Genomic differences/similarities among TuMV lineages and VPg as a potential**
422 **target of selection**

423 Finally, we have explored the molecular changes experienced by the TuMV evolved
424 lineages. Full genomic consensus sequences were determined for 45 evolved lineages.
425 A total of 119 mutational events have been observed, affecting 73 nucleotide positions
426 (supplementary figure S3A and table S1, Supplementary Material online). According to
427 the type of nucleotide substitution involved, 98 were transitions and 21 transversions;
428 regarding their effect on the protein sequence, 30 were synonymous and 89
429 nonsynonymous. Interestingly, some mutations, including 13 nonsynonymous and four
430 synonymous, have been observed multiple times in independent lineages. As discussed
431 in the Materials and Methods section, treating each lineage as an observation and each
432 host genotype as a subpopulation, the average nucleotide diversity within host genotypes,
433 referred only to the 73 polymorphic sites, was $\pi_S = 0.067 \pm 0.007$. On the other hand, the
434 nucleotide diversity for the entire sample was $\pi_T = 0.073 \pm 0.008$. Hence, the estimated
435 interhost genotypes nucleotide diversity was $\delta_{ST} = 0.006 \pm 0.003$, which results in a
436 coefficient of nucleotide differentiation (Nei 1982) of $N_{ST} = 0.081 \pm 0.039$, a value
437 significantly greater than zero ($z = 2.077$, $P = 0.019$). Thus, we conclude that minor yet
438 significant genetic differentiation has been generated among viral lineages replicating in
439 different host genotypes. To assess whether selection played a role in virus genetic
440 differentiation among *A. thaliana* genotypes, we performed a Tajima's *D* test (Tajima
441 1989) and found that it was significantly negative ($D = -2.496$, $P = 0.006$).

442 Next, we sought to characterize the distribution of mutational events along the nine
443 non-overlapping cistrons (supplementary figure S3B, Supplementary Material online).
444 The frequency of mutations per cistron, relative to the length of the corresponding cistron,

445 was fitted to the logistic regression model shown in equation 5 of the Materials and
446 Methods. Highly significant differences exist (χ^2 = 143.206, 9 d.f., P < 0.001), yet
447 entirely due to the ~11-fold larger mutation frequency observed in the *VPg* cistron relative
448 to the rest of the genome (supplementary figure S3B, Supplementary Material online).
449 Notice that all mutations observed in *VPg* are nonsynonymous and that all lineages except
450 *eds8-1*/L3 carry at least one mutation in this cistron (supplementary figure S3A and table
451 S1, Supplementary Material online).

452 Convergent nonsynonymous mutations are, *a priori*, good candidates for adaptive
453 mutations. Mutation CP/V148I appears in two lineages evolved in *cpr5-2*
454 (supplementary figure S3A and supplementary table S1, Supplementary Material online),
455 mutation CP/S70N appear in lineage *dcl2 dcl4*/L1 and three lineages evolved in *jin1*
456 (supplementary figure S3A and table S1, Supplementary Material online), and two
457 different mutations affecting CP amino acid 112, CP/D112G (lineage *npr1-1*/L4) and
458 CP/D112A (lineage *i4g2*/L5), that result in a similar replacement of side chains. In
459 addition, 10 nonsynonymous mutations in *VPg* are shared by several lineages
460 (supplementary figure S3A and table S1, Supplementary Material online). Out of these
461 10 cases, three seem particularly promising candidates. They all are located in a narrow
462 region of VPg (residues 113, 115 and 118) (supplementary figure S3A and table S1,
463 Supplementary Material online). Firstly, mutations G6237A, G6237C, A6238G, and
464 U6239A all affect the same codon, resulting in amino acid replacements D113N, D113H,
465 D113G, and D113E, respectively. The expected functional effect of these mutations was
466 evaluated using the algorithms implemented in SNAP2 (Hecht, Bromberg and Rost
467 2015). D113E is predicted to be functionally neutral (score = –76, accuracy = 87%),
468 while D113G (32, 66%), D113H (43, 71%) and D113N (30, 66%) are predicted to have
469 a functional effect. Among these three amino acid replacements, D113G, which shows
470 an intermediate value of the functional effect score, is particularly interesting, involving
471 a very strong change in the side radical from a long negatively charged one to a small
472 nonpolar one. This replacement has been observed in lineages *cpr5-2*/L3, *cpr5-2*/L4,
473 *cpr5*-2/L5, *npr1-1*/L1, *npr1-1*/L5, *hsp90-1*/L2, and WT/P. All four mutant hosts where
474 these lineages evolved belong to the phenogroup G4 (figure 1). Secondly, mutations
475 A6243C, A6243G and A6244G also affect the same codon, resulting in amino acid
476 replacements N115H (–42, 72%), N115E (–66, 82%) and N115S (–30, 61%). All three
477 are predicted to be functionally neutral. Thirdly, mutations U6252 and G6253A affect

478 the same codon and result in amino acid replacements R118C (–10, 53%) and R118H
479 (56, 78%), respectively. The most frequent among these two replacements is R118H,
480 which represents a conservative change among large positively charged side radicals, yet
481 has a strong expected functional effect. R118H has been observed in four out of five
482 TuMV lineages evolved in *jin1* (phenotypic group G5 in figure 1). D113G thus represents
483 a case of convergent evolution non-specific of the local host genotype (*i.e.*, a candidate
484 for a generalist mutation) while R118H represents a mutation highly specific of the local
485 host genotype (*i.e.*, a specialist mutation).
486 To further characterize the possible adaptive value of mutations VPg/D113G and
487 VPg/R118H, we created by site-directed mutagenesis the two mutated versions of *VPg*
488 and cloned them into the ancestral p35STunos infectious clone (Chen et al. 2003).
489 Viruses were recovered from these clones and three disease-related phenotypic traits (*e.g.*,
490 *AUDPS*, *I* and *SS*) evaluated both in the WT plants as well as in their corresponding local
491 hosts (*i.e.*, *cpr5-2* for mutant VPg/D113G and *jin1* for mutant VPg/R118H). Figure 9
492 summarizes the results of these experiments. Regarding VPg/D113G (figure 9A - C), the
493 results strongly depend on the combination of phenotypic trait and plant genotype: it
494 shows a significant negative effect in both *AUDPS* (–14.2%; $z = 1.687$, $P = 0.046$) and *I*
495 (–23.9%; $z = 6.344$, $P < 0.001$) but a positive one in *SS* (16.7%; $z = 2.597$, $P = 0.005$)
496 when evaluated in WT plants. However, it shows a largely negative yet not significant
497 effect in *cpr5-2* plants for *AUDPS* (–60.8%; $z = 1.055$, $P = 0.146$) and significant negative
498 effects both for *I* (–60.1%; $z = 10.614$, $P < 0.001$) and *SS* (–46.7%; $z = 3.228$, $P = 0.001$).
499 These negative effects on the local host genotype *cpr5-2* do not support a possible
500 beneficial effect of this mutation by itself. Regarding the possible adaptive effect of the
501 amino acid replacement VPg/R118H (figure 9D - F), the results are also dependent on the
502 combination of phenotypic trait and host plant genotype. When tested in WT plants, the
503 mutation has significant negative effects both for *AUDPS* (–11.4%; $z = 2.29$, $P = 0.013$)
504 and *I* (–9.0%; $z = 2.749$, $P = 0.003$) but a significantly positive effect for *SS* (43.6%; $z =$
505 7.881, $P < 0.001$). In contrast, in the local host *jin1* the effect was positive in all three
506 traits; significant for *AUDPS* (9.2%; $z = 1.683$, $P = 0.046$) and *SS* (68.2%; $z = 12.447$, P
507 < 0.001) but not for *I* (0.3%; $z = 1.342$, $P = 0.090$). These results support the conclusion
508 that the amino acid replacement VPg/R118H has an overall beneficial effect on the local
509 host genotype *jin1* in which it was selected.
510

511

512 # Discussion

513 **Interaction between TuMV and different plant defenses**

514 In this study we have explored the effect that mutations in different host disease signaling
515 pathways or in recessive resistance *S* genes have in the outcome of virus evolution. Our
516 experimental pathosystem consisted of two well studied components: the model plant *A.*
517 *thaliana* and one virus with high prevalence in natural populations on this host, TuMV
518 (Pagán et al. 2010). Among the disease signaling pathways, we have studied genes
519 involved in the SA, JA/ET and RNA-silencing; among the *S* genes, we have included heat
520 shock proteins, transcription factors and components of the translation machinery. In a
521 preliminary set of experiments, we compared phenotypic responses of different plant
522 mutant genotypes to TuMV infection. Our results were, in some cases, at odds with the
523 expected phenotypes for *R* genes, the most unexpected one being the result for *jin1*.
524 *JASMONATE INSENSITIVE 1* (*JIN1*) is a negative regulator of SA-mediated defense
525 responses, henceforth the mutant *jin1* has a constitutive expression of SAR (Laurie-Berry
526 et al. 2006). Surprisingly, it turns out to be the most sensitive genotype to TuMV
527 infection, showing enhanced symptoms of disease (phenogroup G5). This may reflect
528 the fact that the function of genes involved in disease signaling have been mostly defined
529 in terms of plant interactions with biotrophic and necrotrophic bacteria and fungi but
530 rarely in response to virus infections. Mutant genotypes only affecting components of
531 the ISR pathway (*e.g.*, *coi1-4*) behaved as WT plants in response to TuMV infection, thus
532 confirming previous reports that ISR was inefficient against viral infections (Ton et al.
533 2002; Loebenstein 2009; Pieterse et al. 2009). However, mutant *eds8-1*, which avoids
534 ISR but enhances SAR (Love et al. 2007) turned out to be among the most resistant
535 genotypes to infection (phenogroup G1). In agreement with our findings of a variable
536 response to SA-signaling, Singh et al. (2004) discussed examples of viruses in which SA-
537 dependent responses to viral infection were strongly dependent on the virus used in the
538 experiments. Replication of viruses such as alfalfa mosaic virus, potato virus X or turnip
539 vein clearing virus was inhibited by treating plants with SA, while cucumber mosaic virus
540 appeared unaffected by the treatment and accumulated to normal levels.

541 The RNA-silencing pathway is considered as the main plant defense against viral
542 pathogens (Voinnet 2001), with *DICER LIKE 2* (*DCL2*) and *DICER LIKE 4* (*DCL4*)
543 encoding for the two dicer enzymes responsible for generating the 22- and 21-nucleotides
544 long antiviral siRNAs. Therefore, the double mutant *dcl2 dcl4* was expected to show a

545 highly sensitive response to TuMV infection. However, infected plants were, overall,
546 hardly distinguishable from the WT plants in their response to infection (supplementary
547 figures S2 and S3, Supplementary Material online). A possible explanation would be the
548 strong suppressor activity of HC-Pro in the WT plants that effectively counteracts the
549 defense mechanisms (Kaschau and Carrington 2001). Indeed, it has been shown that the
550 outcome of the interplay between plant RNA-silencing response and potyviruses is
551 mostly driven by the efficiency of the viral HC-Pro suppression activity (Lin et al. 2007;
552 Torres-Barceló et al. 2008).

553 The response to infection of genotypes carrying mutations in *S* genes was, in
554 general, consistent with the *a priori* expectation. Mutant *i4g2* belonging to phenogroup
555 G1, that is defective for the eIF4(iso)E factor, shows a strong resistance to TuMV
556 infection (Nicaise et al. 2007; Charron et al. 2008). Likewise, in agreement with previous
557 descriptions, mutant *dbp2* belonging to phenogroup G3, shows enhanced resistance to
558 potyvirus infection (Castelló et al. 2011). Finally, mutant *hsp90-1* turns out to show a
559 response to TuMV infection equivalent to the WT plants, likely due to its functional
560 redundancy with other heat-shock proteins present in the cell that may be coopted by the
561 virus to assist in their expected functions (Verchot 2012).

562 An interesting finding was the effect of the knocked down expression of gene
563 $P58^{IPK}$ on TuMV replication. In mammals $P58^{IPK}$, a tetratricopeptide repeat (TPR)-
564 containing protein, is recruited by viruses (*e.g.*, influenza A virus) to inhibit interferon
565 activation and cell death mediated by the dsRNA-activated protein kinase (PKR), thus
566 favoring viral spread (Goodman et al. 2011). In our experiments, $p58^{IPK}$ mutant plants
567 showed enhanced tolerance to infection by showing weaker symptoms than WT plants
568 despite the same level of virus accumulation (supplementary figure S1, Supplementary
569 Material online). In sharp contrast, Bilgin et al. (2003) found that infection of $p58^{IPK}$-
570 silenced *Nicotiana benthamiana* and *A. thaliana* plants infected with tobacco mosaic
571 virus led to death. Given the limited information available from other plant-virus
572 pathosystems, additional experiments would be necessary to shed light into the potential
573 antiviral role in plants of this gene.

574 An important conclusion of our study is that the generally accepted model of plant
575 defense signaling pathways, mainly based in experiments done with bacteria and fungi,
576 may not describe well the interaction between plants and viruses. By contrast, recessive
577 *S* gene-based resistance seems to better explain differences in susceptibility to infection

578 and viral accumulation, thus being a more promising target for future development of
579 resistant plants.
580
581 **Evolution of specialist and generalist viral strategies depends on the host genotype**
582 The GFG and MA models of host-virus interaction modes represent the two ends of a
583 continuum of possible outcomes (Agrawal and Lively 2002). The most relevant
584 difference between both models regards the expected genetic heterogeneity in both host
585 and virus populations. With a pure GFG interaction, the susceptible host types are
586 expected to disappear and the resistant types to dominate the population. *Vice versa*, the
587 most virulent virus allele would become fixed in the population at the cost of mild alleles.
588 However, constitutive activation of defenses is known to be costly for *A. thaliana* [*e.g.*,
589 SA-related defense responses pay a fitness cost in absence of pathogens (Traw et al.
590 2007)] and high virulence usually comes with a cost in terms of pathogen's transmission
591 (Acevedo et al. 2019). Hence, a pure GFG strategy seems unlikely to be achieved. By
592 contrast, with a pure MA interaction, negative frequency-dependent selection emerges,
593 such that rare *A. thaliana* resistance alleles have advantage and, as a result, a genetic
594 polymorphism shall be maintained (Schmid-Hempel 2011). Natural populations of *A.*
595 *thaliana* contain considerable amount of genetic variability for tolerance (Pagán et al.
596 2008) and for immunity-related genes (Todesco et al. 2010, van de Weyer et al. 2019),
597 thus suggesting that an arms race between pathogens and plants is still ongoing but it is
598 unclear whether it may result in pure GFG or MA interactions or lie somewhere in
599 between. Evolution experiments with such different pathosystems as TuMV-*A. thaliana*
600 (González et al. 2019), *Octosporea bayeri-Daphnia magna* (Altermatt and Ebert 2008)
601 and *Serratia marcescens*-*Caenorhabditis elegans* (Gibson et al. 2020; White et al. 2020)
602 have produced congruent results: parasites exposed to heterogeneous host populations
603 evolved significantly lower virulence than parasites exposed to homogeneous host
604 populations. However, a significant difference exists among pathosystems: while viruses
605 exposed to genetically heterogenous host populations evolved as no-cost generalists,
606 evolution of generalism in more complex parasites was constrained by a fitness tradeoff,
607 as expected for the jack-of-all trades hypothesis (Bedhomme et al. 2015).
608 Our results, as well as those by Hillung et al. (2014) and González et al. (2019)
609 have shown the evolution of significantly nested bipartite infection networks, a finding
610 compatible with the existence of a combination of specialist and generalist viruses and of
611 more permissive and resistant host genotypes. Indeed, these studies have also shown that

612 more permissive hosts selected for more specialized viruses while more resistant hosts
613 selected for more generalist viruses, here matching the predictions of the GFG model.
614 Our observation of small yet significant modularity in the infection network was easily
615 explained by convergent evolution of TuMV lineages evolved in the same host genotype.
616 However, it has been recently shown in a long-term survey of the prevalence of different
617 plant viruses in different hosts and habitats that nestedness and modularity in host-
618 pathogen infection networks is possible due to the spatially patched distribution of
619 habitats and temporal successions of plant species (Valverde et al. 2020): small spatial
620 scales create modularity that coexist with global nestedness. This pattern may change
621 spatially and temporally but remains stable over long evolutionary timescales.
622
623 **Role of natural selection**
624 We found evidence of significant genetic differentiation among TuMV lineages evolved
625 in different plant genotypes. To test whether these differences were driven by selection
626 we performed a Tajima's *D* test (Tajima 1989). The resulting negative *D* value was
627 significant, which is compatible with the action of purifying selection, the presence of
628 slightly deleterious mutations segregating in the populations or fast population
629 expansions (Yang 2006). How to distinguish between these explanations? In
630 independent fast expanding populations, many new mutations may be generated and
631 rising in frequency in each population, thus being observed as singletons, mutations
632 present in only one of the many coexisting genomes in each evolving TuMV lineage.
633 Singletons inflate the number of segregating sites and thus cause $D < 0$. Indeed, this is
634 the case here: 56 out of the 73 observed variable sites are singletons, thus the observed
635 pattern of molecular diversity among lineages evolved in the same host genotype and in
636 different host genotypes is likely to be due to the fast expansions of viral populations.
637 However, we have found additional evidence supporting the action of positive
638 selection: the existence of a number of convergent nonsynonymous mutations arising in
639 independent lineages. Some of these nonsynonymous mutations evolved in the same host
640 genotype but others arise in different host genotypes (supplementary table S1 and figure
641 S3, Supplementary Material online). Yet without a clear association with the particular
642 signaling pathway, mutated *S* gene (table 1) or the phenogroup they belong to (figure 1).
643 Interestingly, most of these convergent mutations happened in the *VPg* cistron, which
644 also turns out to be the most variable one. VPg plays many essential roles in genome
645 transcription (it is linked to the 5'-end of the viral genome and provides the hydroxyl

646 group that primes the synthesis of the complementary strains by the viral RdRp),
647 translation [directly interacts with the eukaryotic initiation factors eIF(iso)4E and
648 eIF(iso)4G] and interacts with all other viral proteins (Bosque et al. 2014) and some of
649 the host cell proteins (Martínez et al. 2016; Martínez et al. 2020). Indeed, in previous
650 evolution experiments with potyviruses, *VPg* has also been shown to be an important
651 target of selection. For example, Agudelo-Romero et al. (2008) found that a single amino
652 acid replacement in VPg was enough to largely increase TEV *I*, *SS*, *VL* and in *A. thaliana*.
653 Similarly, Gallois et al. (2010) found that *A. thaliana* plants with knock-out mutations in
654 the *eIF(iso)4E*, *eIF(iso)4G1* and *eIF(iso)4G2* genes were resistant to TuMV infection.
655 Two mutations in the VPg (E116Q and N163Y) were enough to overcome this resistance
656 and return to the original infection phenotype, though yeast-two hybrid assays showed
657 that none of these mutations affected the binding of VPg with eIF(iso)4E (Gallois et al.
658 2010). As a final example, one of the most extensively used resistance genes against
659 potato virus Y in commercial pepper cultivars is pvr^2, which has many different alleles
660 (Nicaise et al. 2007; Charron et al. 2008). The pvr^2 locus encodes for the eIF4E factor
661 which, as mentioned above, physically interacts with VPg. Interestingly, all the
662 resistance-breaking viral isolates found so far contain mutations in the *VPg* cistron
663 (Duprat et al. 2002; Moury et al. 2004; Ayme et al. 2006).

664 One of the two mutations identified by Gallois et al. (2010), VPg/E116Q, affects
665 the same protein domain as mutations D113G and R118H identified and characterized
666 here. The mechanism by which these two mutations may confer a selective advantage to
667 TuMV lineages cannot be inferred from our studies. We found VPg/D113G in several
668 lineages evolved in different host genotypes, while VPg/R118H was only found in the
669 *jin1* lineage. Our assays failed to find a beneficial effect of VPg/D113G but confirmed
670 the host-specific beneficial effect of VPg/R118H in *jin1* plants.

671 Unfortunately, we have not been able to exhaustively test all observed mutations
672 and, hence, several others may still be potential candidates for adaptive mutations. More
673 specifically, we have not tested possible epistatic interactions among mutations fixed in
674 the same genome. For example, three lineages carry more than one nonsynonymous
675 mutation in *VPg* (*dbp2*/L1 D113N and N115E, *eds8-1*/L1 N115H and E116G, and *eds8-*
676 *1*/L4 H33Y, D113N and K121E) (supplementary table S1, Supplementary Material
677 online). In addition, 25 lineages carry additional nonsynonymous mutations in at least
678 one other cistron besides *VPg* (supplementary table S1, Supplementary Material online).
679 Two lineages have up to four additional nonsynonymous mutations: (*i*) *cpr5-2*/L5 that

680 has mutations in proteins P3/T318M-T326S, CI/V148I and VPg/D113G and (*ii*) *dcl2*
681 *dcl4*/L1 that has mutations in proteins P1/V54A, CI/I378V, VPg/R118C, and CP/S70N.
682 Therefore, plenty of opportunities for epistatic effects both within and among cistrons
683 exist, and will be evaluated in future experiments.
684
685
686 # Methods
687 **Plants, virus and growth conditions**
688 A collection of 21 different *A. thaliana* mutants of the Col-0 accession were used for this
689 study (table 1). In all experiments described below, plants were all maintained in a BSL2
690 climatic chamber under a photoperiod of 8 h light (LED tubes at PAR 90 - 100 μmol/m^2/s)
691 at 24 ºC and 16 h dark at 20 ºC and 40% relative humidity.
692 Prior to the inoculation experiments, we created a large stock of TuMV infectious
693 saps. Saps were obtained from TuMV-infected *N. benthamiana* Domin plants inoculated
694 with the infectious plasmid p35STunos that contains a cDNA of TuMV genome
695 (GeneBank accession AF530055.2) under the control of the cauliflower mosaic virus 35S
696 promoter and the *nos* terminator (Chen et al. 2003) as described elsewhere (González et
697 al. 2019; Corrêa et al. 2020). This TuMV sequence variant corresponds to the YC5 isolate
698 from calla lily (*Zantesdeschia* sp.) (Chen et al. 2003). After the plants showed symptoms
699 of infection, they were pooled and frozen with liquid N_2. This frozen plant tissue was
700 homogenized into a fine powder using a Mixer Mill MM400 (Retsch GmbH, Haan,
701 Germany). For inoculations, the 0.01 g of powder was diluted in 1 mL inoculation buffer
702 (50 mM phosphate buffer pH 7.0, 3% PEG6000, 10% Carborundum) and 5 μL of the
703 inoculum was gently rubbed into the plant leaves. Plants were all inoculated when they
704 reached the growth stage 3.5 in the Boyes et al. (2001) scale. This synchronization
705 ensures that all hosts were at the same phenological stage when inoculated.
706
707 **Phenotyping infected plants**
708 Five different disease-related traits were measured for each infected plant for 18 dpi. (*i*)
709 Change in dry weight (ΔDW) of the aerial part of infected plants, with a precision of 10
710 mg, relative to the corresponding noninfected controls (supplementary figure S1A;
711 Supplementary Material online). (*ii*) Severity of symptoms (*SS*) evaluated in a semi-
712 quantitative discrete scale (figure 1B in Corrêa et al. 2020) ranging from zero for

713 asymptomatic infections to four for plants showing a generalized necrosis and wilting
714 (supplementary figure S1B, Supplementary Material online). (*iii*) The area under the
715 disease progress stairs (*AUDPS*), that summarizes the speed at which the disease
716 progresses in a group of plants (supplementary figure S1C, Supplementary Material
717 online) and takes values in the range (0, *N*) where *N* is the total number of plants included
718 in the assay (Simko and Piepho 2012). (*iv*) Infectivity measured as the number of
719 symptomatic plants out of the number of inoculated plants at 18 dpi (*I*; supplementary
720 figure S1D, Supplementary Material online). And (*v*) viral load (*VL*) measured by
721 absolute RT-qPCR as the number of viral genomes per ng of total RNA in the plant as
722 described below (supplementary figure S1E, Supplementary Material online).
723 The five traits were not all independent but showed some significant pairwise
724 positive correlations: *AUDPS* with *I* ($r = 0.772$, 19 df, $P < 0.001$), *SS* ($r = 0.807$, 19 df, P
725 < 0.001) and *VL* ($r = 0.472$, 19 df, $P = 0.031$), and *I* with *SS* ($r = 0.746$, 19 df, $P < 0.001$).
726 Basically, ΔDW is orthogonal with the other four traits while the other four traits showed
727 some degree of association. Interesting is the case of *VL*, which was only (weakly)
728 correlated to *AUDPS*.
729
730 **Experimental evolution**
731 Five TuMV lineages were evolved during 12 consecutive serial passages in each one of
732 the nine selected mutant genotypes. In addition, two lineages were also evolved in WT
733 plants. To begin the evolution experiment, ten 21-days old *A. thaliana* plants per lineage
734 and genotype were inoculated as described above with the virus stock previously created
735 from infected tissue of *N. benthamiana* plants inoculated with an TuMV isolate YC5
736 infectious clone. The next passages were made by harvesting the symptomatic plants at
737 14 dpi, preparing the infectious sap as described above and inoculating it to a new healthy
738 population of ten plants. These saps were 1/10 diluted with inoculation buffer and used
739 to inoculate the next batch of plants.
740
741 **Total RNA extractions**
742 Tissue from each pool of infected plants per lineage, genotype and serial passage was
743 collected, frozen with liquid N_2 and preserved at −80 ºC until it was homogenized into
744 fine powder using a Mixer Mill MM400. Next, an aliquot of approximately 100 mg of
745 grounded tissue was used for total RNA (RNAt) extraction. The RNAt was extracted
746 with the Agilent Plant RNA isolation Mini kit (Agilent Technologies, Santa Clara CA,

747 USA). Aliquots of RNAt per each sample were separated and their concentration adjusted
748 at approximately 50 ng/µL to estimate viral accumulation by RT-qPCR (see below).
749
750 **Quantification of *VL***
751 *VL* of each plant sample per lineage, genotype and selected passage was quantified by
752 absolute real-time quantitative RT-PCR (RT-qPCR) using standard curves and the
753 primers TuMV F117 forward (5'-CAATACGTGCGAGAGAAGCACAC-3') and F118
754 reverse (5'-TAACCCCTTAACGCCAAGTAAG-3') that amplify a 173 nucleotides
755 fragment from the *CP* cistron of TuMV genome, as previously described (Corrêa et al.
756 2020). Briefly, standard curves were constructed using ten serial dilutions of the TuMV
757 genome, that was synthesized by *in vitro* transcription as detailed previously (Cervera et
758 al. 2018), in RNAt extract obtained from healthy *A. thaliana* plants used as control in the
759 experiments. Amplification reactions were run in a 20 µL volume using the GoTaq 1-
760 Step RT-qPCR System (Promega, Madison WI, USA) and the recommended
761 manufacturer's instructions as described in Cervera et al. (2018) in an ABI StepOne Plus
762 Real-time PCR System (Applied Biosystems, Foster City CA, USA). The cycling
763 conditions consisted in: an RT phase of 5 min at 42 ºC and 10 min at 95 ºC followed by
764 a PCR stage consisting in 40 cycles of 5 s at 95 ºC and 34 s at 60 ºC; and the final melt
765 curve profile that consisted in 15 s at 95 ºC, 1 min at 60ºC and 15 s at 95 ºC. Negative
766 controls consisted of healthy RNAt plant extract (mock-inoculated non infected control)
767 and water. For each sample, three technical replicates were quantified. The results were
768 analysed using the StepOne software 2.2.2 (Applied Biosystems).
769
770 **TuMV genome amplifications and sequencing**
771 Evolved viral genomes of the passage 12 from each lineage and genotype were amplified
772 by high-fidelity RT-PCR using the AccuScript Hi-Fi (Agilent Technologies) reverse
773 transcriptase and Phusion DNA polymerase (Thermo Scientific, Waltham MA, USA)
774 following the manufacturer's instructions. Each complete TuMV genome was amplified
775 into three overlapping amplicons of 3114 (5' fragment R1), 3697 (central region R2) and
776 3287 nucleotides (3' fragment R3) using three primer sets. For RT reactions an aliquot
777 of the corresponding RNAt (150-300 ng) was mixed with 0.25 µM of the 1R-P3 (5'-
778 CGAGTAGTATCTTATAGCACAGCGCTCCGACC-3'), 2R-NIa (5'-
779 TGTCTGGAATCGGTAGCAAATGTAGCTGAGTTGTG-3') or 3R-polyAR (5'-
780 TTTTTTTTTTTTTTTTTTTTGTCCCTTGCATCATATCAAATG-3') primer to

781 synthesize the R1, R2 or R3 cDNA fragment, respectively, that were denatured 5 min at
782 65 ºC and cooled on ice. Then a mix containing AccuScript Hi-Fi 1× Buffer, 1 mM of
783 dNTPs, 8 mM of DTT, 4U of Ribolock RNase inhibitor (Thermo Scientific) and 0.5 µL
784 of AccuScript Hi-Fi (Agilent Technologies) was added up to a 10 µL volume. RT
785 reactions consisted of 90 min at 42 ºC to synthesize the cDNA followed by an incubation
786 of 5 min at 70 ºC to inactivate the enzyme. PCR reactions were performed in a 50 µL
787 volume containing a mix of 1× Phusion Buffer, 0.4 µM of dNTPs, 0.2 µM of each primer,
788 0.5-1 µL of DMSO, 2U of Phusion DNA polymerase (Thermo Scientific) and 1 µL of the
789 corresponding RT reaction. R1 fragment was amplified using the primer set 1F-5UTR
790 (5'-GCAAACGCAGACCTTTCGAAGCACTCAAGC-3') and 1R-P3 and the following
791 PCR conditions: an initial denaturation of 30 s at 98 ºC, 3 cycles of 10 s at 98 ºC, 20 s at
792 67 ºC and 2 min at 72 ºC, 3 cycles of 10 s at 98 ºC, 20 s at 65 ºC and 2 min at 72 ºC, and
793 32 cycles of 10 s at 98 ºC, 20 s at 63 ºC and 2 min at 72 ºC, followed by a final extension
794 step of 5 min at 72 ºC. Fragments R2 and R3 were amplified with primer set 2F-P3 (5'-
795 TGGGAGCTTGCGGATGGTGGATACACAATTC-3') and 2R-NIa or 3F-NIa (5'-
796 CTCGTTATATGGAGTCGGTTTCGGACCACTCATAT-3') and 3R-polyAR,
797 respectively and a PCR with the same denaturation and extension steps than the fragment
798 R1 but different amplification steps: a stage consisting of 15 cycles of 10 s at 98 ºC, 20 s
799 at 67 ºC and 2 min at 72 ºC followed by 23 cycles of 10 s at 98 ºC, 20 s at 65 ºC and 2
800 min at 72 ºC for fragment R2. For fragment R3 15 cycles of 10 s at 98 ºC, 20 s at 67 ºC
801 and 2 min at 72 ºC followed by 23 cycles of 10 s at 98 ºC, 20 s at 65 ºC and 2 min at 72
802 ºC were used. PCR products were purified with the MSB Spin PCRapace Kit (Stratec
803 Molecular, Coronado CA, USA) and then Sanger-sequenced. Full-length consensus viral
804 sequences were obtained assembling the sequences of the three amplified products by
805 using the Genious R9.0.2 program.
806
807 **Construction of TuMV *VPg* mutants**
808 Mutants in the *VPg* cistron were synthesized by long inverse site-directed PCR
809 mutagenesis using the QuickChange II XL Kit (Stratagene, San Diego CA, USA)
810 following the manufacturer's instructions and using p35STunos as template. The
811 mutagenic primers were designed following the manufacturer's recommendations.
812 Primer set cpr5-2F (5'-GGAGGATGAGTTGGGTCCAAATGAAATACGTGT-3') and
813 cpr5-2R (5'- ACACGTATTTCATTTGGACCCAACTCATCCTCC-3') was used to

814 generate the A6238G (D113G) mutant while primer set jin1F (5'-
815 GGATCCAAATGAAATACATGTGAATAAGACAATTC-3') and jin1R (5'-
816 GAATTGTCTTATTCACATGTATTTCATTTGGATCC-3') to obtain the G6253A
817 (R118H) mutant. PCR protocol consisted of: a denaturation step of 2 min at 95 ºC
818 followed by 20 cycles of 20 s at 95 ºC, 10 s at 60 ºC and 7 min at 68 ºC, and a final
819 extension of 5 min at 68 ºC. After *Dpn*I digestion and transformation of electrocompetent
820 *Escherichia coli* DH5α, the presence of the desired mutations in the infectious clone and
821 the absence of undesired nucleotide changes was confirmed by sequencing.
822
823 **Bioassays of the TuMV *VPg* mutants**
824 The infectivity, viability and pathogenicity of both VPg TuMV mutants was confirmed
825 by performing three independent bioassays. In each experiment, three batches of 24 *A.*
826 *thaliana* plants (three-weeks old) of genotypes WT, *cpr5-2* and *jin1* were inoculated with
827 the WT TuMV plasmid clone used as reference and with plasmids of each of the TuMV
828 *VPg* mutants. Another two batches of 24 plants from *cpr5-2* or *jin1* genotypes were
829 inoculated with the WT TuMV plasmid and the corresponding plasmid of the TuMV VPg
830 mutant, the VPg/D113G mutant in the case of the *cpr5-2* genotype and that of
831 VPg/R118H for the *jin1* genotype. *A. thaliana* plants were inoculated by abrasion of
832 three leaves applying equal amounts of the plasmid inoculum, a total of approximately 7
833 µg per plant distributed in 3 µL per leaf. Plasmids were previously purified by Midiprep
834 using the NucleoBond Xtra Midi Kit (Macherey-Nagel, Düren, Germany) and
835 resuspended in distilled water. Number of infected plants and symptom intensity were
836 collected every day until 15 dpi.
837
838 **Statistical analyses**
839 All statistical analyses described hereafter were performed with SPSS version 27 software
840 (IBM, Armonk, NY), unless otherwise indicated.
841 *AUDPS* (figure 2), *I* (figure 3) and *VL* (figure 4) data were fitted all together to a
842 multivariate analysis of covariance (MANCOVA) model with plant genotype (*G*) as main
843 factor, independent evolution lineages (*L*) were nested within *G* and passage (*t*) was
844 introduced in the model as a covariable. The full model equation thus reads:
845 $\mathbf{P}_{ijk}(t) \sim \mathbf{\Pi} + t + G_i + L(G)_{ij} + (t \times G)_i + [t \times L(G)]_{ij} + \boldsymbol{\varepsilon}_{ijk}$, (Eq. 1)

where $\mathbf{P}_{ijk}(t) = (AUDPS_{ijk}, I_{ijk}, VL_{ijk})^{T}$ is the vector of phenotypic traits observed at time t, for an individual infected plant k of evolutionary lineage j of genotype i, $\mathbf{\Pi}$ represents the vector of phenotypic grand mean values and $\boldsymbol{\varepsilon}_{ijk}$ stands for the vector of errors assumed to be Gaussian distributed at every t. The significance of each factor, the covariable and their interactions was evaluated using the Wilks' Λ method. The magnitude of the effects was evaluated using the η_P^2 statistic (proportion of total variability in the traits vector attributable to each factor in the model; conventionally, values of $\eta_P^2 \geq 0.15$ are considered as large effects).

Rates of phenotypic evolution for *AUDPS* and *I* were estimated by fitting the time series data to a first-order autoregressive integrated moving-average, ARIMA(1,0,0), model (Elena and Sanjuán, 2005; González et al. 2019). The model equation fitted has the form:

$$Y_t - \rho_1 Y_{t-1} = Y_0 + \nu_Y t + \varepsilon_t, \qquad \text{(Eq. 2)}$$

where Y_k represents the variable being analyzed at passage k, ρ_1 measures the degree of self-similarity in the time-series data (correlation between values at passages t and $t-1$), ε_t represents the sampling error at passage t, and ν_Y represents the linear dependency of variable Y with passage number, that is, the rate of phenotypic evolution. The rates of phenotypic evolution were further analyzed in the context of the mutated defense signaling pathway genes or recessive resistances. Rates of phenotypic evolution were fitted to a multivariate analysis of variance (MANOVA) in which the plant genotype, serving as host for the evolutionary lineages (G), was nested within the factor being analyzed (*i.e.*, the type of selection –hard *vs* soft-, the mode of resistance –signaling pathways *vs S*-, and whether mutations affected the SA-dependent, the JA/ET-dependent signaling or the RNA-silencing pathways; table 1). The full model equation now thus reads:

$$\mathbf{V}_{ijk} \sim \mathbf{\Phi} + X_{ij} + X(G)_{ij} + \boldsymbol{\varepsilon}_{ijk}, \qquad \text{(Eq. 3)}$$

where $\mathbf{V}_{ijk} = (\nu_{AUDPS}, \nu_I)^{T}$ is the vector of rates of phenotypic evolution observed for the k replicate lineage evolved in plant genotype j within the factor X_i, $\mathbf{\Phi}$ represents the vector of grand mean values and $\boldsymbol{\varepsilon}_{ijk}$ the vector of Gaussian errors.

Analysis of infection network

877 The first statistical approach consisted in fitting a logistic regression model to the
878 presence/absence of symptoms data using GLM techniques with a Binomial probability
879 distribution and a probit link function. The model equation reads as follows:

$$\text{probit}[f_{ijkl}(t)] \sim \phi + t + TG_i + t \times TG_i + LG_j + t \times LG_j + L(LG)_{jk} + t \times L(LG)_{jk} + (TG \times LG)_{ij} + t \times (TG \times LG)_{ij} + [TG \times L(LG)]_{ijk} + t \times [TG \times L(LG)]_{ijk} + \varepsilon_{ijkl}, \quad \text{(Eq. 4)}$$

883 where $f_{ijkl}(t)$ is the frequency of symptomatic plants t dpi in the test host genotype TG, for
884 the viral lineage L evolved in the local host genotype LG. TG and LG were considered
885 as orthogonal factors, L was nested within LG and dpi (t) was treated as a covariable. The
886 model includes all main factors, their corresponding nested and factorial interactions as
887 well as their interactions with the covariable. ϕ represents the grand mean of the probit-
888 transformed f values, and ε the error term assumed to be Binomial. The significance of
889 the different factors was evaluated using likelihood-ratio tests (LRT).

890 For the second statistical approach, the binary infection matrix was analyzed using
891 tools borrowed from the field of network biology to explore whether they show random
892 associations between viral lineages and host genotypes, one-to-one associations,
893 nestedness indicative of a GFG type of interaction, or modularity (Weitz et al. 2013). The
894 statistical properties of the infection matrix were evaluated using the R package
895 "bipartite" version 2.15 (Dormann and Strauss 2008) in R version 4.0.0 (R Core Team
896 2020) under RStudio 1.2.1335. Four different summary statistics were evaluated: T
897 nestedness (Bascompte et al. 2003), Q modularity (Newman 2006) and the d' species-
898 level (or Kullback-Leibler divergence) and H_2' network-level (or two-dimensional
899 normalized Shannon entropy) specialization indexes (Blüthgen et al. 2006). H_2' ranges
900 between zero and one for extreme generalists and specialists, respectively. Statistical
901 significance of these statistics was evaluated using Bascompte et al. (2003) null model.
902

903 **The statistics of molecular evolution**

904 Treating each lineage as an observation and each host genotype as a subpopulation, we
905 evaluated the average nucleotide diversity within host genotypes, π_S, the nucleotide
906 diversity for the entire sample, π_T, the interhost genotypes nucleotide diversity, δ_{ST}, and
907 the estimate of the proportion of interhost genotypes nucleotide diversity, known as
908 coefficient of nucleotide differentiation (Nei 1982), $N_{ST} = \delta_{ST}/\pi_T$. Standard deviations of
909 estimates were inferred from 1000 bootstrap samples. All these computations were done

using MEGA 11 (Tamura et al. 2021) and the lowest-BIC nucleotide substitution model Kimura 2-parameters (Kimura 1980). Tajima's *D* test of selection (Tajima 1989) and its statistical significance were evaluated using DnaSP6 (Rozas et al. 2017).

The frequency of mutations (*m*) per cistron (*C*), relative to the length of the corresponding cistron, was fitted to the following logistic regression model using GLM techniques with a Binomial probability distribution and a probit link function:

$$\text{probit}(m_i) \sim \mu + C_i + \varepsilon_i \qquad \text{(Eq. 5)}$$

where μ is the average genomic mutation frequency and *i* refers to the 10 cistrons in the main ORF.

The functional effect of mutations found in the *VPg* cistron were evaluated *in silico* using the SNAP2 algorithm (Hecht et al. 2015). The algorithm provides a functionality score and an accuracy index for each possible mutation affecting a coding sequence. Negative score values mean functional neutrality while positive values should be taken as indications of functional effects; the larger the values, the stronger the effect.

Supplementary Material

Supplementary data are available *at Molecular Biology and Evolution* online.

Acknowledgments

We thank Francisca de la Iglesia and Paula Agudo for excellent technical assistance and the rest of the EvolSysVir lab members for fruitful discussions. This work was supported by grants BFU2015-65037-P and PID2019-103998GB-I00 (Agencia Estatal de Investigación - FEDER) and PROMETEU2019/012 (Generalitat Valenciana) to S.F.E.

References

Acevedo MA, Dillemuth FP, Flick AJ, Faldyn MJ, Elderd BD. 2019. Virulence-driven trade-offs in disease transmission: a meta-analysis. *Evolution* 73:636-647.

Agrawal A, Lively CM. 2002. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evol Ecol Res*. 4:79-90.

942 Agudelo-Romero P, Carbonell P, Pérez-Amador MA, Elena SF. 2008. Virus adaptation
943 by manipulation of host's gene expression. *PLoS ONE* 3:e2397.
944 Altermatt F, Ebert D. 2008. Genetic diversity of Daphnia magna populations enhances
945 resistance to parasites. *Ecol Lett*. 11:918-928.
946 Altizer S, Kaitala V, Laakso J, Ruokolainen L. 2006. Seasonality and the dynamics of
947 infectious diseases. *Ecol Lett*. 9:467-484.
948 Anttila J, Kaitala V, Laakso J, Ruokolainen J. 2015. Environmental variation generates
949 environmental opportunist pathogen outbreaks. *PLoS ONE* 10:e0145511.
950 Ayme V, Souche S, Caranta C, Jacquemond M, Chadoeuf J, Palloix A, Moury B. 2006.
951 Different mutations in the genome-linked protein VPg of potato virus Y confer
952 virulence on the *pvr2*^{3} resistance in pepper. *Mol Plant Microbe Interact.* **19**: 557-
953 563.
954 Bascompte J, Jordano P, Melián CJ, Olesen JM. 2003. The nested assembly of plant-
955 animal mutualistic networks. *Proc Natl Acad Sci USA.* 100:9389-9387.
956 Bedhomme S, Hillung J, Elena SF. 2015. Emerging viruses: why are not jacks of all
957 trades? *Curr Opin Virol.* 10:1-6.
958 Berngruber T, Lion S, Gandon S. 2015. Spatial structure, transmission modes and the
959 evolution of viral exploitation strategies. *PLoS Pathog*. 11:e1004810.
960 Bilgin DD, Liu Y, Schiff M, Dinesh-Kumar SP. 2003. *P58*^{IPK}, a plant ortholog of double-
961 stranded RNA-dependent protein kinase PKR inhibitor, functions in viral
962 pathogenesis. *Dev Cell.* 4:651-661.
963 Blüthgen N, Menzel F, Blüthgen N. 2006. Measuring specialization in species interaction
964 networks. *BMC Ecol*. 6:9.
965 Boots M, Mealor M. 2007. Local interactions select for lower pathogen infectivity.
966 *Science* 315:1284-1286.
967 Bosque G, Folch-Fortuny A, Picó J, Ferrer A, Elena SF. 2014. Topology analysis and
968 visualization of *Potyvirus* protein-protein interaction network. *BMC Syst Biol*. 8:129.
969 Bouché N, Lauressergues D, Gasciolli V, Vaucheret J. 2006. An antagonistic function for
970 Arabidopsis DCL2 in development and a new function for DCL4 in generating viral
971 siRNAs. *EMBO J.* 25:3347-3356.
972 Boyes DC, Zayed AM, Ascenzi R, McCaskill MJ, Hoffman NE, Davis KR, Görlach J.
973 2001. Growth stage-based phenotypic analysis of Arabidopsis: a model for high
974 throughput functional genomics in plants. *Plant Cell* 13:1499-1510.

975 Brockhurst M, Buckling A, Rainey P. 2006. Spatial heterogeneity and the stability of
976 host-parasite coexistence. *J Evol Biol*. 19:374-379.
977 Brown JKM, Tellier A. 2011. Plant-parasite coevolution: bridging the gap between
978 genetics and ecology. *Annu Rev Phytopathol*. 49:345-367.
979 Cao H, Bowling SA, Gordon AS, Dong X. 1994. Characterization of an Arabidopsis
980 mutant that is nonresponsive to inducers of systemic acquired resistance. *Plant Cell*
981 6:1583-1592.
982 Carr JP, Lewsey MG, Palukaitis P. 2010. Signaling in induced resistance. *Adv Virus Res*.
983 76:57-121.
984 Castelló MJ, Carrasco JL, Navarrete-Gómez M, Daniel J, Granot D, Vera P. 2011. A plant
985 small polypeptide is a novel component of DNA-binding protein phosphatase 1-
986 mediated resistance to plum pox virus in Arabidopsis. *Plant Physiol*. 157:2206-2215.
987 Cervera H, Ambrós S, Bernet GP, Rodrigo G, Elena SF. 2018. Viral fitness correlates
988 with the magnitude and direction of the perturbation induced in the host's
989 transcriptome: the *Tobacco etch potyvirus*-tobacco case study. *Mol Biol Evol*.
990 35:1599-1615.
991 Chabas H, Lion S, Nicot A, Meaden S, van Houte, S, Moineu S, Wahl LM, Westra ER,
992 Gandon S. 2018. Evolutionary emergence of infectious diseases in heterogeneous
993 host populations. *PLoS Biol*. 16:e2006738.
994 Charron C, Nicolaï, M, Gallois JL, Robaglia C, Moury B, Palloix A, Caranta, C. 2008.
995 Natural variation and functional analyses provide evidence for coevolution between
996 plant eIF4E and potyviral VPg. *Plant J*. 54:56-68.
997 Chen CC, Chao CH, Chen CC, Yeh SD, Tsai HT, Chang CA. 2003. Identification of
998 turnip mosaic virus isolates causing yellow stripe and spot on calla lily. *Plant Dis*.
999 87:901-905.
1000 Corrêa RL, Sanz-Carbonell A, Kogej Z, Müller SY, Ambrós S, López-Gomollón S,
1001 Gómez G, Baulcombe DC, Elena SF. 2020. Viral fitness determines the magnitude
1002 of transcriptomic and epigenomic reprogramming of defense responses in plants. *Mol*
1003 *Biol Evol*. 37:1866-1881.
1004 Cui H, Qiu J. Zhou Y, Bhandari DD, Zhao C, Bautor J, Parker JE. 2018. Antagonism of
1005 transcription factor MYC2 by EDS1/PAD4 complexes bolsters salicylic acid defense
1006 in Arabidopsis effector-triggered immunity. *Mol Plant* 11:1053-1066.
1007 Dormann CF, Gruber B, Fruend J. 2008. Introducing the bipartite package: analyzing
1008 ecological networks. *R News* 8:8-11.

1009 Dormann CF, Strauss R. 2014. A method for detecting modules in quantitative bipartite
1010 networks. *Meth Ecol Evol*. 5:9-98.
1011 Duprat A, Caranta C, Revers F, Menand B, Browning KS, Robaglia C. 2002. The
1012 Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but
1013 required for susceptibility to potyviruses. *Plant J*. 32:927-934.
1014 Elena SF, Sanjuán R. 2005. RNA viruses as complex adaptive systems. *BioSystems*
1015 81:31-41.
1016 Flor HH. 1956. The complementary genetic systems in flax and flax rush. *Adv Genet*.
1017 8:29-54.
1018 Frank SA. 1993. Specificity versus detectable polymorphism in host-parasite genetics.
1019 *Proc R Soc B.* 254:191-197.
1020 Gallois JL, Charron C, Sánchez F, Pagny G, Houvenaghel MC, Moretti A, Ponz F, Revers
1021 F, Caranta C, German-Retana, S. 2010. Single amino acid changes in the turnip
1022 mosaic virus viral genome-linked protein (VPg) confer virulence towards
1023 *Arabidopsis thaliana* mutants knocked out for eukaryotic initiation factors
1024 eIF(iso)4E and eIF(iso)4G. *J Gen Virol*. 91:288-293.
1025 Gandon S. 2004. Evolution of multihost parasites. *Evolution* 58:455-469.
1026 Ganusov VV, Bergstrom CT, Antia R. 2002. Within-host population dynamics and the
1027 evolution of microparasites in a heterogeneous host population. *Evolution* 56:213-
1028 223.
1029 Gibson AK, Baffoe-Bonnie H, Penley MJ, Lin J, Owens R, Khalid A, Morran LT. 2020.
1030 The evolution of parasite host range in heterogeneous host populations. *J Evol Biol*.
1031 33:773-782.
1032 González R, Butković A, Elena SF. 2019. Role of host genetic diversity for susceptibility-
1033 to-infection in the evolution of virulence of a plant virus. *Virus Evol*. 5:vez024.
1034 Goodman AG, Tanner BCW, Chang ST, Esteban M, Katze MG. 2011. Virus infection
1035 rapidly activates the $P58^{IPK}$ pathway, delaying peak kinase activation to enhance viral
1036 replication. *Virology* 417:27-36.
1037 Gupta V, Willits MG, Glazebrook J. 2000. *Arabidopsis thaliana EDS4* contributes to
1038 salicylic acid (SA)-dependent expression of defense responses: evidence for
1039 inhibition of jasmonic acid signaling by SA. *Mol Plant Microbe Interact*. 13:503-
1040 511.
1041 Haraguchi Y, Sasaki A. 2000. The evolution of parasite virulence and transmission rate
1042 in a spatially structured population. *J Theor Biol*. 203:85-96.

1043 Hecht M, Bromberg Y, Rost B. 2015. Better prediction of functional effects for sequence
1044 variants. *BMC Genomics* 16:S1.
1045 Hillung J, Cuevas JM, Valverde S, Elena SF. 2014. Experimental evolution of an
1046 emerging plant virus in host genotypes that differ in their susceptibility to infection.
1047 *Evolution* 68:2467-2480.
1048 Hughes WHO, Boomsma JJ. 2006. Does genetic diversity hinder parasite evolution in
1049 social insect colonies? *J Evol Biol*. 19:132-143.
1050 Kachroo P, Chandra-Shekara AC, Klessig DF. 2006. Plant signal transduction and
1051 defense against viral pathogens. *Adv Virus Res*. 66:161-191.
1052 Kaschau K, Carrington JC. 2001. Long-distance movement and replication maintenance
1053 functions correlate with silencing suppression activity of potyviral HC-Pro. *Virology*
1054 285:71-81.
1055 Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions
1056 through comparative studies of nucleotide sequences. *J Mol Evol*. 16:111-120.
1057 Lalić J, Agudelo-Romero P, Carrasco P, Elena SF. 2010. Adaptation of *Tobacco etch*
1058 *potyvirus* to a susceptible ecotype of *Arabidopsis thaliana* capacitates it for systemic
1059 infection of resistant ecotypes. *Phil Trans R Soc B.* 65:1997-2008.
1060 Laurie-Berry N, Joardar V, Street IH, Kunkel BN. 2006. The *Arabidopsis thaliana*
1061 *JASMONATE INSENSITIVE 1* gene is required for suppression of salicylic acid-
1062 dependent defenses during infection by *Pseudomonas syringae*. *Mol Plant Microbe*
1063 *Interact*. 19:789-800.
1064 Lin SS, Wu HW, Jan FJ, Hou RF, Yeh SD. 2007. Modifications of the helper component-
1065 protease of zucchini yellow mosaic virus for generation of attenuated mutants for
1066 cross protection against severe infection. *Phytopathology* 97:287-296.
1067 Lively CM. 2010. The effect of host genetic diversity on disease spread. *Am Nat*.
1068 175:e149-e152.
1069 Loebenstein G. 2009. Local lesions and induced resistance. *Adv Virus Res*. 75:73-117.
1070 Love AJ, Laval V, Geri C, Laird J, Tomos AD, Hooks MA, Milner JJ. 2007. Components
1071 of Arabidopsis defense- and ethylene-signaling pathways regulate susceptibility to
1072 cauliflower mosaic virus by restricting long-distance movement. *Mol Plant Microbe*
1073 *Interact*. 20:659-670.
1074 Mäkinen K. 2019. Plant susceptibility genes as a source for potyvirus resistance. *Ann*
1075 *Appl Biol*. 173:122-129.

1076 Martínez F, Rodrigo G, Aragonés V, Ruiz M, Lodewijk I, Fernández U, Elena SF, Daròs
1077 JA. 2016. Interaction network of *Tobacco etch potyvirus* NIa protein with the host
1078 proteome during infection. *BMC Genomics* 17:1.
1079 Martínez F, Toft C, Hillung J, Carrasco JL, Giménez-Santamarina S, Yenush L, Rodrigo
1080 G, Elena SF. 2020. A comprehensive physical interaction map between the *Turnip*
1081 *mosaic potyvirus* and *Arabidopsis thaliana* proteomes. *Res. Square*
1082 doi:10.21203/rs.3.rs-149993/v2.
1083 Mine A, Nobori T, Salazar-Rodon MC, Winkelmüller TM, Anver S, Becker D, Tsuda K.
1084 2017. An incoherent feed-forward loop mediates robustness and tunability in a plant
1085 immune network. *EMBO Rep*. 18:464-476.
1086 Moreno-Gámez S, Stephan W, Tellier L. 2013. Effect of disease prevalence and spatial
1087 heterogeneity on polymorphism maintenance in host-parasite interactions. *Plant*
1088 *Pathol*. 62:133-141.
1089 Moury B, Morel C, Johansen E, Guilbaud L, Souche S, Ayme V, Caranta C, Palloix A,
1090 Jacquemond M. 2004. Mutations in potato virus Y genome-linked protein determine
1091 virulence towards recessive resistances in *Capsicum annuum* and *Lycopersicon*
1092 *hirsutum*. *Mol Plant Microbe Interact*. 17:322-329.
1093 Nawrath C, Heck, S, Parinthawong N, Métraux JP. 2002. *EDS5*, an essential component
1094 of salicylic acid-dependent signaling for disease resistance in Arabidopsis, is a
1095 member of the MATE transporter family. *Plant Cell* 14:275-286.
1096 Nawrath C, Métraux JP. 1999. Salicylic acid induction-deficient mutants of Arabidopsis
1097 express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen
1098 inoculation. *Plant Cell* 11:1393-1404.
1099 Nei M. 1982. Evolution of human races at the gene level. In *Human Genetics, part A: the*
1100 *Unfolding Genome* (ed. B. Bonné-Tamir). Alan R. Liss, New York.
1101 Newman MEJ. 2006. Modularity and community structure in networks. *Proc Natl Acad*
1102 *Sci USA.* 103:8577-8582.
1103 Nicaise V, Gallois JL, Chafiai F, Allen LM, Schurdi-Levraud V, Browning KS,
1104 Candresse T, Caranta C, Le Gall O, German-Retana S. 2007. Coordinated and
1105 selective recruitment of eIF4E and eIF4G factors for potyvirus infection in
1106 *Arabidopsis thaliana*. *FEBS Lett*. 581:1041-1046.
1107 Pagán I, Alonso-Blanco C, García-Arenal F. 008. Host responses in life-history traits and
1108 tolerance to virus infection in *Arabidopsis thaliana*. *PLoS Pathog*. 4:e1000124.

1109 Pagán I, Fraile A, Fernández-Fueyo E, Montes N, Alonso-Blanco C, García-Arenal F.
1110 2010. *Arabidopsis thaliana* as a model for the study of plant-virus co-evolution. *Phil*
1111 *Trans R Soc B* 365:1983-1995.
1112 Parrat SR, Numminen E, Laine AL. 2016. Infectious disease dynamics in heterogeneous
1113 landscapes. *Annu Rev Ecol Evol Syst*. 47:283-306.
1114 Pfenning K. 2001. Evolution of pathogen virulence: the role of variation in host
1115 phenotype. *Proc R Soc B.* 268:755-760.
1116 Pieterse CMJ, León-Reyes A, van der Ent S, van Wees SC. 2009. Networking by small-
1117 molecule hormones in plant immunity. *Nat Chem Biol*. 5:308-316.
1118 Pieterse CMJ, van der Does D, Zamioudis C, León-Reyes A, van Wees SCM. 2012.
1119 Hormonal modulation of plant immunity. *Annu Rev Cell Dev Biol*. 28:489-521.
1120 Regoes RR, Nowak MA, Bonhoeffer S. 2000. Evolution of virulence in a heterogeneous
1121 host population. *Evolution* 54:64-71.
1122 Rodríguez D, Torres-Sorando L. 2001. Models of infectious diseases in spatially
1123 heterogeneous environments. Bull Math Biol. 63:547-571.
1124 Rozas J, Ferrer-Mata A, Sánchez-del Barrio JC, Guirao-Rico S, Librado P, Ramos-Osins
1125 SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of
1126 large data sets. *Mol Biol Evol*. 34:3299-3302.
1127 Sallinen S, Norberg A, Susi H, Laine, A.L. 2020. Intraspecific host variation plays a key
1128 role in virus community assembly. *Nat Commun*. 11:5610.
1129 Schmid-Hempel P. 2011. *Evolutionary Parasitology*. Oxford University Press, Oxford:
1130 UK, pp. 259-272.
1131 Schmid-Hempel P, Koella JC. 1994. Variability and its implications for host-parasite
1132 interactions. *Parasitol Today* 10:98-102.
1133 Simko I, Piepho HP. 2012. The area under the disease progress stair: calculation,
1134 advantage, and application. *Phytopathology* 102:381-389.
1135 Singh DP, Moore CA, Gilliland A, Carr JP. 2004. Activation of multiple antiviral defence
1136 mechanisms by salicylic acid. *Mol Plant Pathol*. 5:57-63.
1137 Sofer L, Cabanillas D, Gayral M, Téplier R, Pouzoulet J, Ducousso M, Dufin L, Bréhélin
1138 C, Ziegler-Graff V, Brault V, Revers Frédéric. 2017. Identification of host factors
1139 potentially involved in *RTM*-mediated resistance during potyvirus long distance
1140 movement. *Arch Virol*. 162:1855-1865.
1141 Soosaar JLM, Burch-Smith TM, Dinesh-Kumar SP. 2005. Mechanisms of plant
1142 resistance to viruses. *Nat Rev Microbiol*. 3:789-799.

1143 Tajima F. 1989. Statistical methods for testing the neutral mutation hypothesis by DNA
1144 polymorphism. *Genetics* 123:585-595.
1145 Tamura K, Stecher G, Kumar S. 2021. Molecular evolutionary genetics analysis version
1146 11. *Mol Biol Evol* 38: 3022-3027.
1147 Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, Nomura K, He SY, Howe
1148 GA, Browse J. 2007. JAZ repressor proteins are targets of the SCF(COI1) complex
1149 during jasmonate signalling. *Nature* 448:661-665.
1150 Todesco M, Balasubramanian S, Hu TT, Traw MB, Horton M, Epple P, Kuhns C,
1151 Sureshkumar S, Schwartz C, Lanz C, Laitinen RAE, Huang Y, Chory J, Lipka V,
1152 Borevitz JO, Dangl JL, Bergelson J, Nordborg M, Weigel D. 2010. Natural allelic
1153 variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature*
1154 465:632-636.
1155 Ton J, van Pelt JA, van Loon LC, Pieterse CMJ. 2002. Differential effectiveness of
1156 salicylate-dependent and jasmonate/ethylene-dependent induced resistance in
1157 Arabidopsis. *Mol Plant Microbe Interact*. 15:27-34.
1158 Torres-Barceló C, Martín S, Daròs JA, Elena SF. 2008. From hypo- to hypersuppression:
1159 effect of amino acid substitutions on the RNA-silencing suppressor activity of the
1160 tobacco etch potyvirus HC-Pro. *Genetics* 180:1039-1049.
1161 Travisano M, Mongold, JA, Bennett, AF, Lenski RE. 1995. Experimental test of the roles
1162 of adaptation, chance, and history in evolution. *Science* 267:87-90.
1163 Traw MB, Kniskern JM, Bergelson J. 2007. SAR increases fitness of *Arabidopsis*
1164 *thaliana* in the presence of natural bacterial pathogens. *Evolution* 61:2444-2449.
1165 Valverde S, Vidiella B, Montáñez R, Fraile A, Sacristán S, García-Arenal F. 2020.
1166 Coexistence of nestedness and modularity in host-pathogen infection networks. *Nat*
1167 *Ecol Evol*. 4:568-577.
1168 Van de Weyer AL, Monteiro F, Furzer OJ, Nishimura MT, Cevik V, Witek K, Jones JDG,
1169 Dangl JL, Weigel D, Bemm F. 2019. A species-wide inventory of NLR genes and
1170 alleles in *Arabidopsis thaliana*. *Cell* 178:1260-1272.
1171 Van Wees SCM, de Swart EAM, van Pelt JA, van Loon, LC, Pieterse CMJ. 2000.
1172 Enhancement of induced disease resistance by simultaneous activation of salicylate-
1173 and jasmonate-dependent defense pathways in *Arabidopsis thaliana*. *Proc Natl Acad*
1174 *Sci USA*. 97:8711-8716.

1175 Verchot J. 2012. Cellular chaperones and folding enzymes are vital contributors to
1176 membrane bound replication and movement complexes during plant RNA virus
1177 infection. *Front Plant Sci*. 3:271.
1178 Voinnet O. 2001. RNA silencing as a plant immune system against viruses. *Trends Genet*.
1179 17:449-459.
1180 Weitz JS, Poisot T, Meyer JR, Flores CO, Valverde S, Sullivan MB, Hochberg ME. 2013.
1181 Phage-bacteria infection networks. *Trends Microbiol*. 21:82-91.
1182 White PS, Choi A, Pandey R, Menezes A, Penley M, Gibson AK, de Roode J, Morran L.
1183 2020. Host heterogeneity mitigates virulence evolution. *Biol Lett.* 16:20190744.
1184 Yang Z. 2006. *Computational Molecular Evolution*. Oxford University Press. Oxford:
1185 UK, pp 262-264.
1186 Yang Z, Huang Y, Yang J, Yao S, Zhao K, Wang D, Qin Q, Bian Z, Li Y, Lan Y, Zhou
1187 T, Wang H, Liu C, Wang W, Qi Y, Xu Z, Li Y. 2020. Jasmonate signaling enhances
1188 RNA silencing and antiviral defense in rice. *Cell Host Microbe* 28:89-103.
1189 Yates A, Antia R, Regoes RR. 2006. How do pathogen evolution and host heterogeneity
1190 interact in disease emergence? *Proc R Soc B.* 273: 3075-3083.
1191 Zhou JM, Zhang Y. 2020. Plant immunity: danger perception and signaling. *Cell*
1192 181:978-989.

1193 **Figures and Tables**

1194

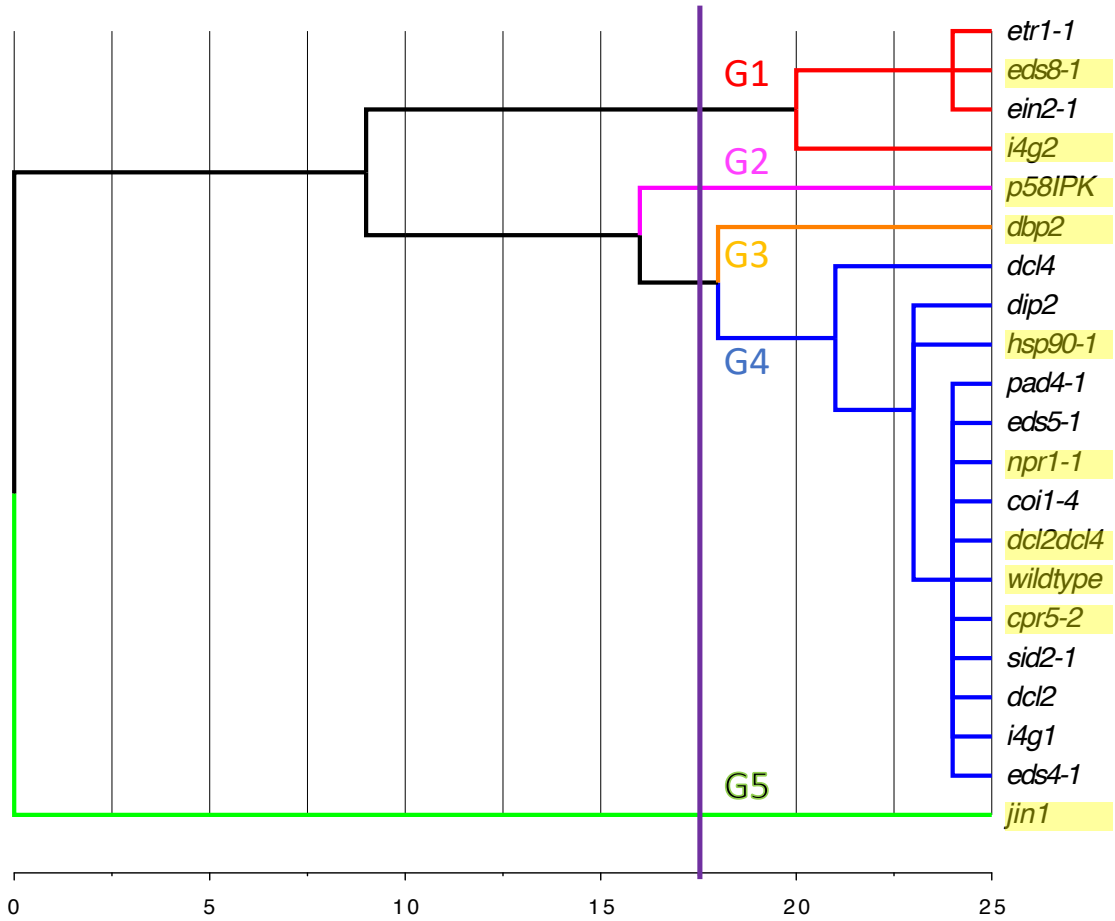

FIG. 1. Nearest-neighbor clustering, based on the squared Euclidean distance, of the 21 mutant genotypes of *A. thaliana* according to their phenotypic similarity in response to TuMV infection. Genotypes selected as hosts for the evolution experiments are highlighted in yellow.

1195

1196

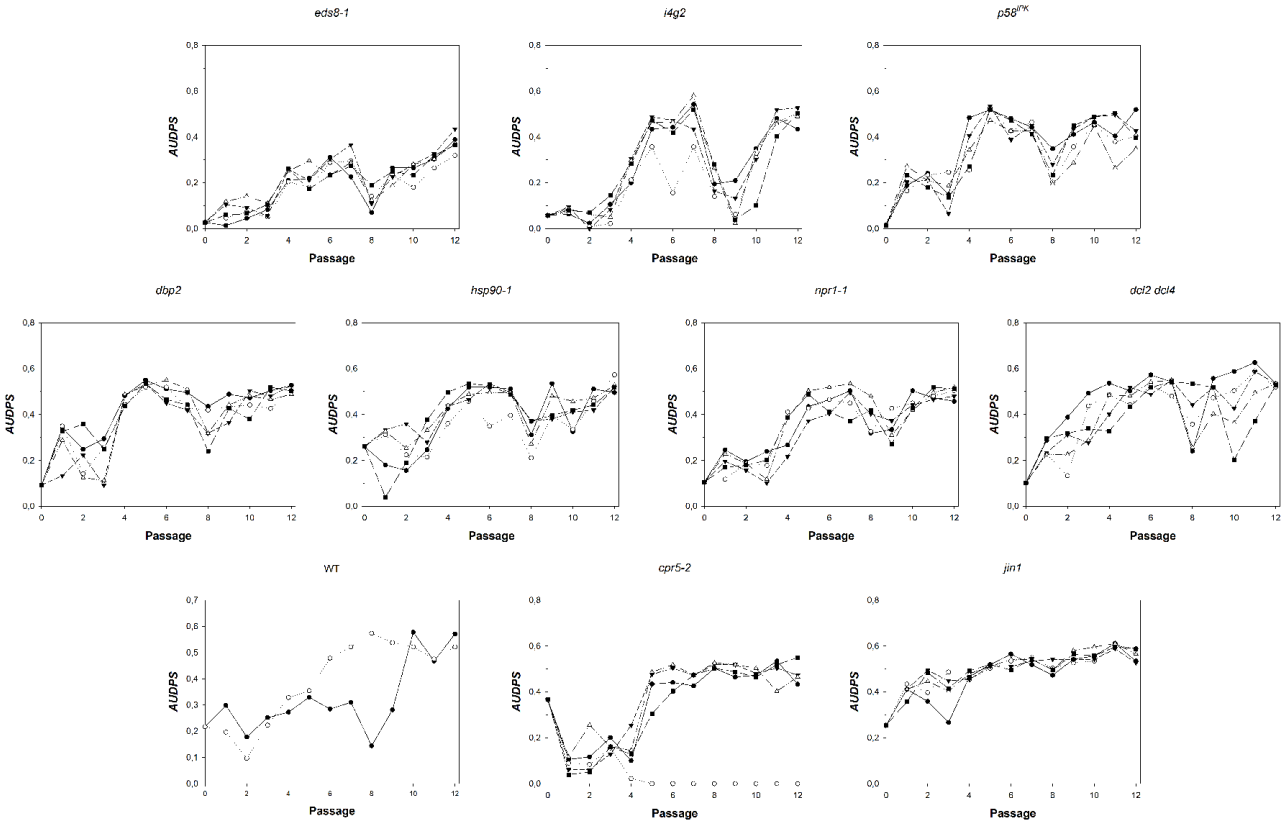

FIG. 2. Evolution of disease progression (*AUDPS*) along the serial passages of experimental evolution on each different host genotype. Different symbols and lines represent the independent evolutionary lineages. Panels are arranged from the most resistant genotype (*eds8-1*) to the most sensitive one (*jin1*) according to the groups defined in figure 1.

1197

1198

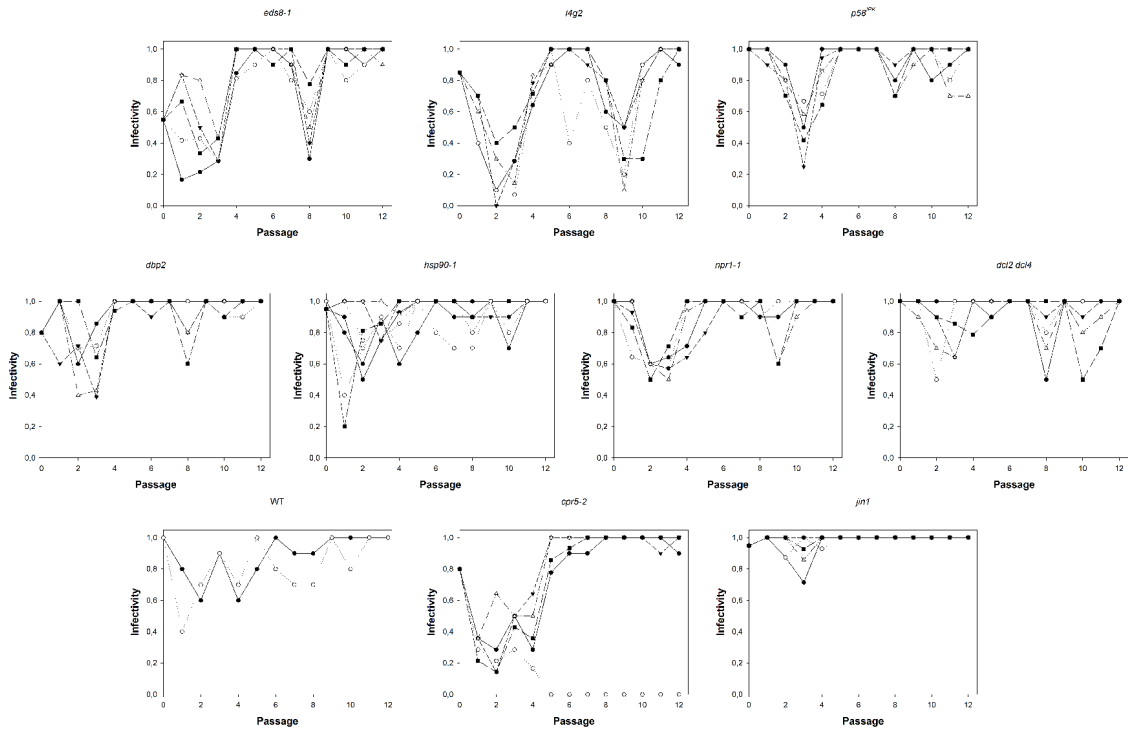

FIG. 3. Evolution of infectivity (*I*) along the serial passages of experimental evolution on each different host genotype. Different symbols and lines represent the independent evolutionary lineages. Panels are arranged from the most resistant genotype (*eds8-1*) to the most sensitive one (*jin1*) according to the groups defined in figure 1.

1199

1200

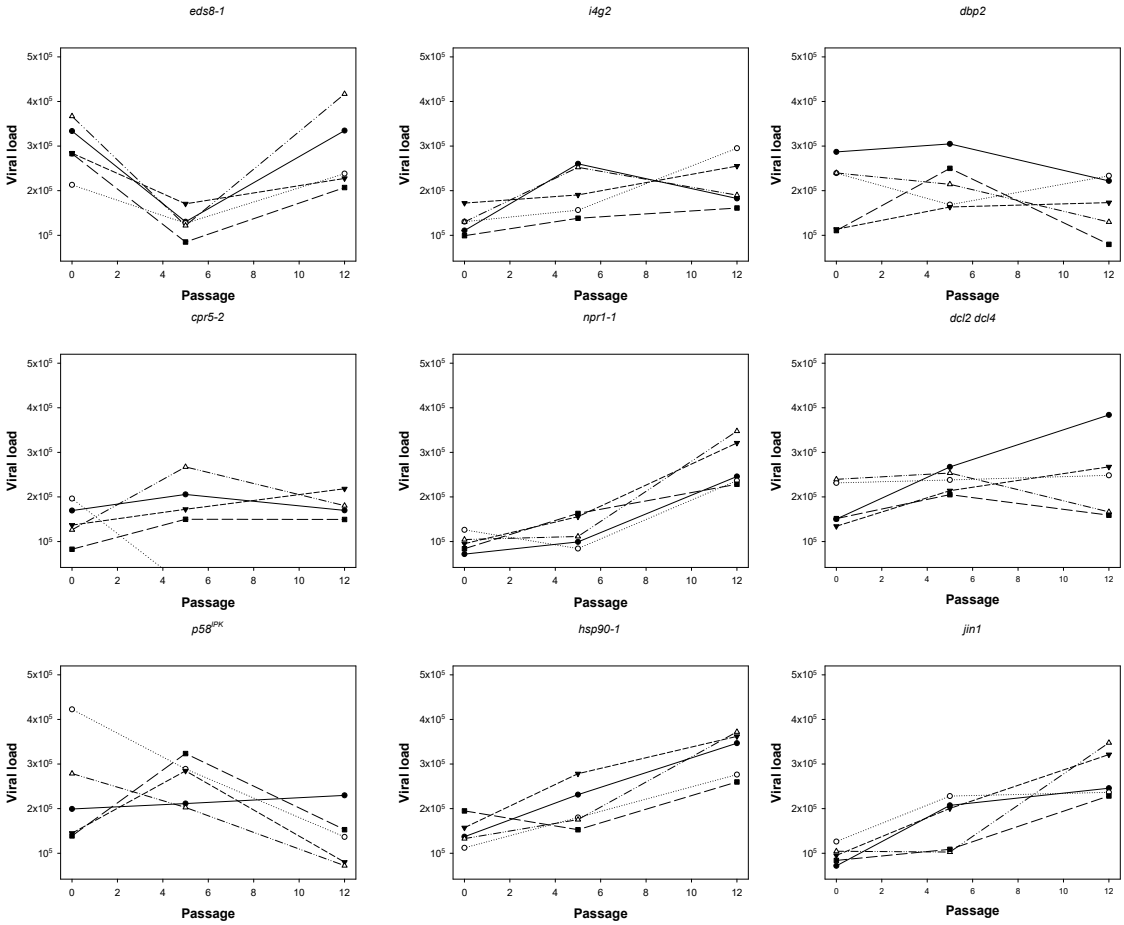

FIG. 4. Evolution of viral load (*VL*) along the serial passages of experimental evolution on each different host genotype. Different symbols and lines represent the independent evolutionary lineages. Panels are arranged from the most resistant genotype (*eds8-1*) to the most sensitive one (*jin1*) according to the groups defined in figure 1. TuMV *VL* was not evaluated for WT plants along the evolution experiment.

1201

1202

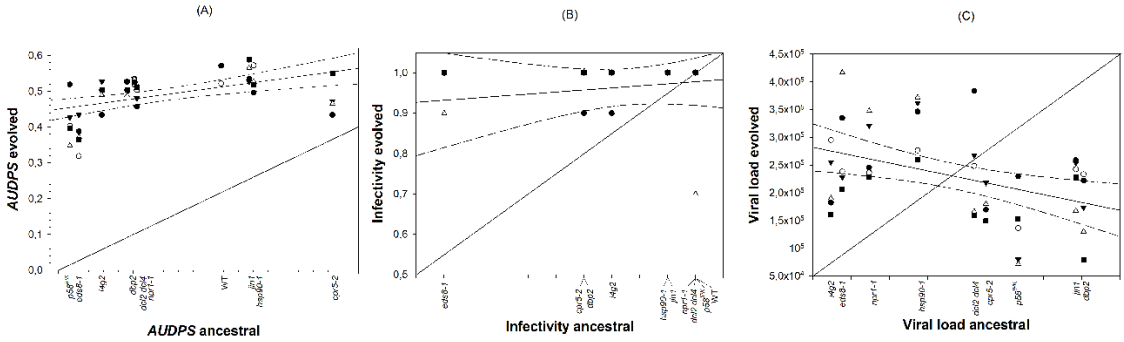

FIG. 5. Test of the contribution of historical contingency to the observed pattern of adaptation of TuMV to the different *A. thaliana* genotypes. (A) Evolved *vs* ancestral values for *AUDPS*, (B) for *I*, and (C) for *VL*. The diagonal lines represent the null hypothesis of historical differences being fully preserved despite adaptation. The solid lines represent the linear regression of the data (dashed lines represent the 95% CI for the regression lines).

1203

1204

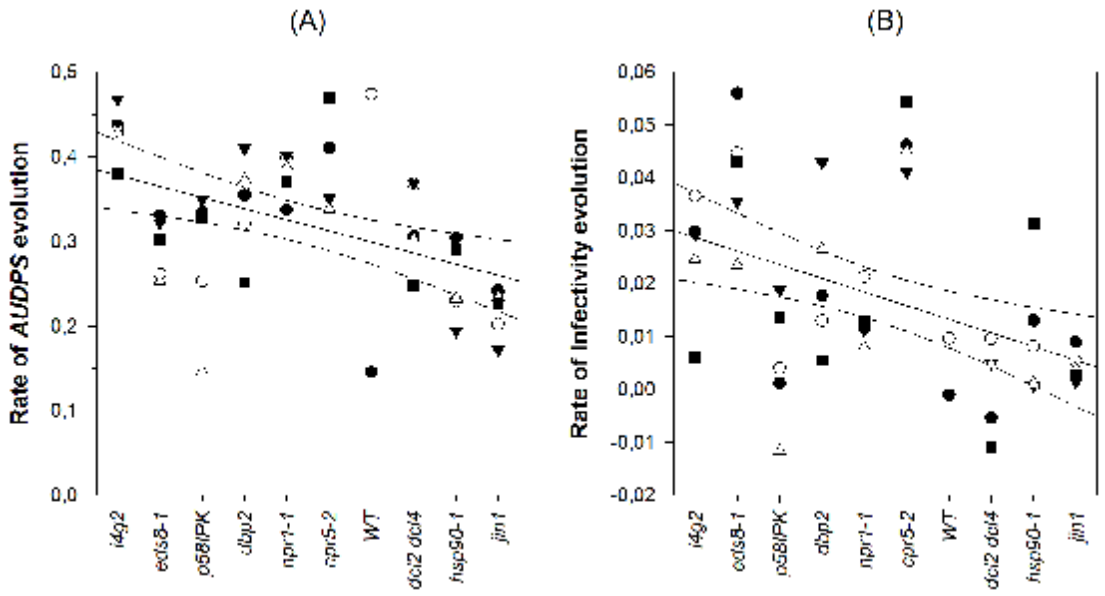

FIG. 6. Estimated rates of evolution for *AUDPS* (A) and *I* (B) obtained from the fitting of an ARIMA(1,0,0) model (equation 2) to the data shown in figure 2 and figure 3, respectively. Different symbols represent different independent lineages evolved on the corresponding host genotype. Solid lines represent the linear regression of the data; dashed lines the 95% confidence intervals for the regression lines. In both cases, *A. thaliana* genotypes are ranked from the less to the more permissive to TuMV infection according to figure 1.

1205

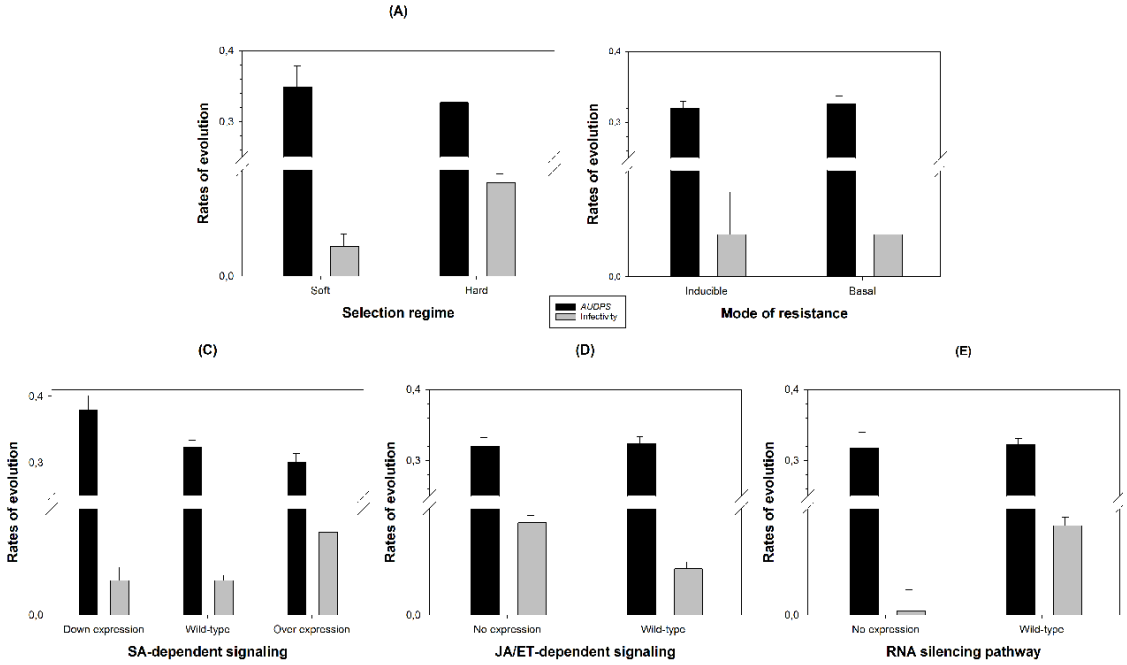

FIG. 7. The rates of evolution for *AUDPS* and *I* differ among possible mechanisms of resistance to infection. Rates of evolution are expressed in log-scale to better visualize the slowest rates of *I* evolution. Data were fitted to the MANOVA model defined by equation. 3. Error bars represent ±1 SEM.

1206

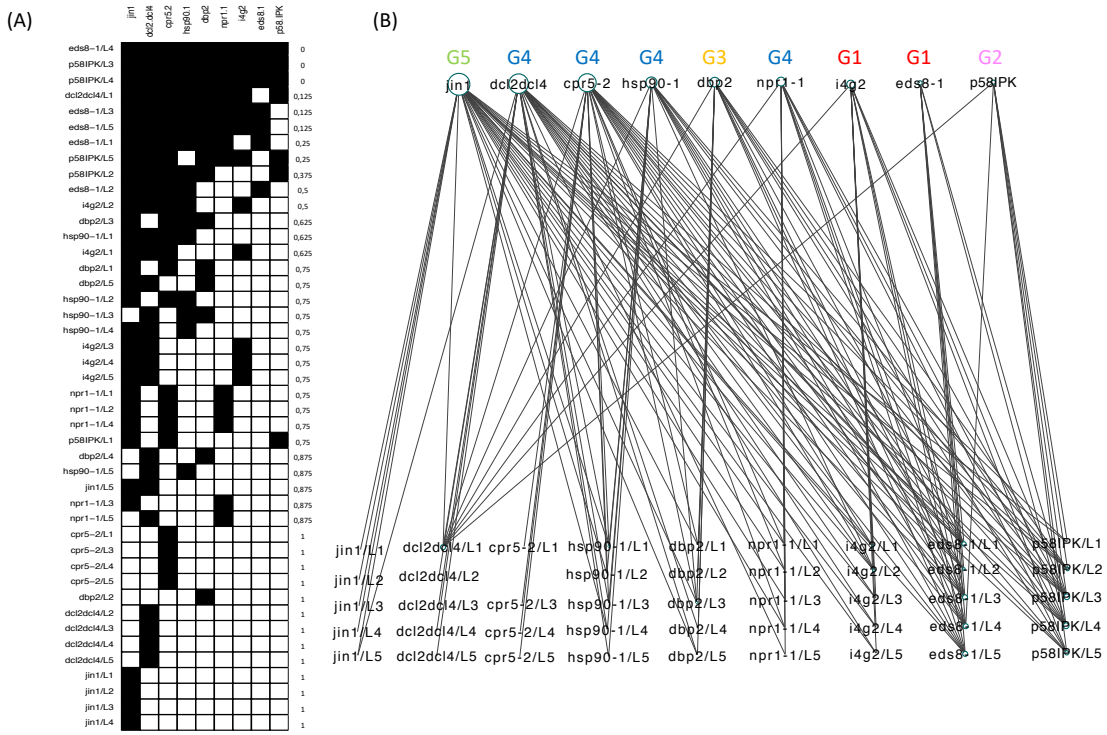

FIG. 8. Analysis of the full cross-infection matrix. (A) Packed matrix that highlights its nested structure, compatible with a gene-for-gene infection model. Black squares represent cases in which the *AUDPS* values were equal or greater than those observed for the corresponding lineage in its local host. Last column shows the species-level specialization index *d'* (*d'* = 0 means most generalist pathogen, *d'* = 1 means most specialist pathogen). (B) Bipartite host genotype - viral lineage infection network. The size of the nodes is proportional to their degree. The phenotypic groups defined in figure 1 are indicated above the host genotype.

1207

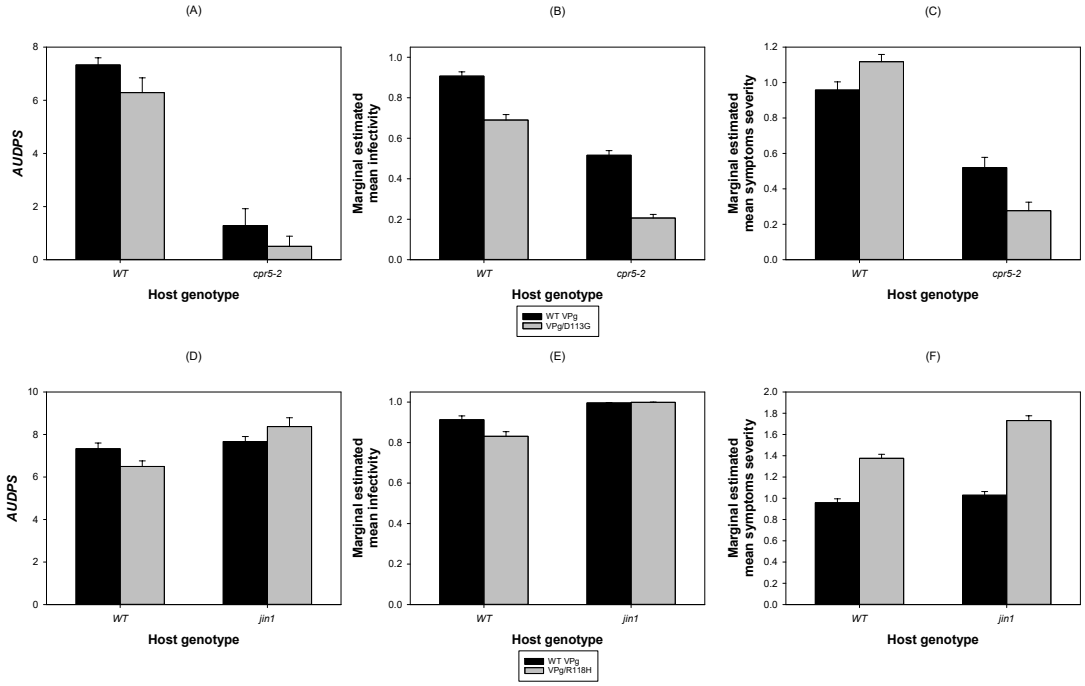

Fig. 9. Phenotypic disease-related traits (*AUDPS*, *I* and *SS*) of TuMV VPg/D113G relative to WT VPg. (A) – (C) Evaluated in the WT *A. thaliana* and the *cpr5-2* mutant genotype. (D) – (F) Evaluated in the WT *A. thaliana* and the *jin1* mutant genotype. Error bars represent ±1 SEM.

1208

Table 1. Different arabidopsis genotypes used in this study. Highlighted in gray are those used for the evolution experiments.

Genotype	Gene name	Affected pathway	Expected phenotype relative to wildtype plants	Reference
Wildtype				
coi1-4	*CORONATINE INSENSITIVE 1* (AT2G39940)	Repression of JA-responsive genes	No ISR, no effect on virus	Thines *et al.* (2007)
cpr5-2	*CONSTITUTIVE EXPRESSOR OF PR GENES 5* (AT5G64930)	Membrane protein, negative regulator of pathogen-dependent SA signaling	More resistant, constitutive SAR	Love *et al.* (2007)
dbp2	*DNA-BINDING PROTEIN PHOSPHATASE 2*	Transcriptional regulation of gene expression in potyvirus-infected plants	More resistant to potyvirus infection	Castelló *et al.* (2011)
dcl2	*DICER-LIKE 2* (AT3G03300)	Partial loss of RNA-silencing	No effect, siRNAs produced by DCL4	Bouché *et al.* (2006)
dcl4	*DICER-LIKE 4* (AT5G20320)	Partial loss of RNA-silencing	No effect, siRNA produced by DCL2	Bouché *et al.* (2006)
dcl2 dcl4	Double mutant *dcl2 dcl4*	Complete loss of RNA-silencing	More susceptible, no siRNA production	Bouché *et al.* (2006)
dip2	*DBP-INTERACTING PROTEIN 2* (AT5G03210)	Transcriptional regulation of gene expression in potyvirus-infected plants	More susceptible	Castelló *et al.* (2011)
eds4-1	*ENHANCED DISEASE SUSCEPTIBILITY 4* (AT5G51200)	Loss of SA-dependent signaling	More susceptible, no SAR	Gupta *et al.* (2000)
eds5-1	*ENHANCED DISEASE SUSCEPTIBILITY 5* (AT4G39030)	Lipase-like protein, positive regulator of pathogen-dependent SA signaling	More susceptible, no SAR	Nawrath *et al.* (2002)

eds8-1	*ENHANCED DISEASE SUSCEPTIBILITY 8*	Reduced expression of plant defensin genes, reduced ISR	More resistant, enhanced SAR	Love *et al*. (2007)
ein2-1	*ETHYLENE INSENSITIVE 2* (AT5G03280)	MAPK, ET signaling intermediate, negative regulator SA-dependent signaling	More resistant, enhanced SAR	Love *et al*. (2007)
etr1-1	*ETHYLENE RESPONSE 1* (AT1G66340)	ET receptor, negative regulator SA-dependent signaling	More resistant, enhanced SAR	Love *et al*. (2007)
hsp90-1	*HEAT SHOCK PROTEIN 1* (AT5G52640)	Recessive r gene, required for membrane-bound replication complexes; protein folding	More resistant, missing component for viral replication	Verchot (2012)
i4g1	*EUKARYOTIC TRANSLATION INITIATION FACTOR (ISO) 4G 1* (AT3G60240)	Recessive *r* gene, initiation of viral RNA translation	More resistant, missing component for viral gene expression	Nicaise *et al*. (2007)
i4g2	*EUKARYOTIC TRANSLATION INITIATION FACTOR (ISO) 4G 2*	Recessive *r* gene, initiation of viral RNA translation	More resistant, missing component for viral gene expression	Nicaise *et al*. (2007)
jin1	*JASMONATE INSENSITIVE 1* (AT1G32640)	Loss of JA signaling; negative regulator of SA-dependent signaling	More resistant, enhanced SAR	Laurie-Berry *et al*. (2006)
npr1-1	*NONEXPRESSER OF PR GENES 1* (AT1G64280)	Ankyrin-repeat protein required for *PR-1* activation	More susceptible, no SAR, no ISR	Cao *et al*. (1994)
p58^{IPK}	*HOMOLOG OF MAMMALIAN P58^{IPK}* (AT5G03160)	Constitutive activation of PKR	More resistant, strong apoptosis-mediated HR	Bilgin *et al*. (2003)

pad4-1	*PHYTOALEXIN DEFICIENT 4* (AT3G52430)	Lipase-like protein, positive regulator pathogen-dependent SA signaling	More susceptible, no SAR	Cui *et al*. (2018)
sid2-1	*SA INDUCTION DEFICIENT 2* (AT1G74710)	Isochorismate synthase, required for SA bioshynthesis	More susceptible, no SAR	Nawrath and Métraux (1999)

ET - ethylene; HR – hypersensitive response; ISR - induced systemic resistance; JA - jasmonic acid; MAPK – mitogen-activated protein kinase; PKR - protein kinase RNA-activated; SA - salicylic acid; SAR - systemic acquired resistance

1209

Table 2. Results of the multivariate analysis of variance (MANCOVA) for the three phenotypic traits evaluated along the course of experimental evolution. The different factors are defined in equation 1.

Source of variation	Wilk's Λ	F	df_1, df_2	P	η_P^2	$1-\beta$
Π	0.004	3377.616	3, 42	< 0.001	0.996	1.000
t	0.265	38.799	3, 42	< 0.001	0.735	1.000
G	0.093	6.467	24, 122.414	< 0.001	0.547	1.000
$t\times G$	0.155	4.608	24, 122.414	< 0.001	0.463	1.000
$L(G)$	0.529	0.278	108, 126.692	1.000	0.191	0.373
$t\times L(G)$	0.466	0.351	105, 126.677	1.000	0.225	0.487

η_P^2: magnitude of the effect; $1-\beta$: power of the test.

1210

1211

Table 3. Results of the probit regression testing for the specificity of adaptation. The different model factors are defined in equation. 4.

Source of variation	LRT	df	P	η_P^2	$1-\beta$
ϕ	21224.493	1	< 0.001	0.959	1.000
t	25995.056	1	< 0.001	0.815	1.000
TG	295.160	8	< 0.001	0.569	1.000
$t\times TG$	568.212	8	< 0.001	0.010	0.981
LG	20.545	8	0.008	0.288	0.661
$t\times LG$	30.200	8	< 0.001	0.005	0.804
$L(LG)$	207.257	35	< 0.001	0.385	1.000
$t\times L(LG)$	235.073	35	< 0.001	0.016	0.984
$TG\times LG$	258.888	64	< 0.001	0.314	1.000
$t\times TG\times LG$	365.512	64	< 0.001	0.018	0.961
$TG\times L(LG)$	788.661	280	< 0.001	0.048	0.999
$t\times TG\times L(LG)$	1033.649	280	< 0.001	0.042	0.993

LRT: likelihood ratio test distributed as a χ^2; η_P^2: magnitude of the effect; $1-\beta$: power of the test.

1212