1 **Defects in plant immunity modulate the rates and**

2 **patterns of RNA virus evolution**

3

4 **Rebeca Navarro[1], Silvia Ambrós[1], Fernando Martínez[1], Beilei Wu[1,$], José L.**

5 **Carrasco[1], and Santiago F. Elena[1,2,*]**

6

7 [1]Instituto de Biología Integrativa de Sistemas (CSIC – Universitat de València), Paterna,

8 46182 València, Spain.

9 [2]The Santa Fe Institute, Santa Fe NM87501, USA.

10

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

12 [$]Permanent address: Institute of Plant Protection, Chinese Academy of Agricultural

13 Sciences, Haidian, 100193 Beijing, China.

14

15 *Keywords*: experimental evolution; generalism; plant immunity; plant-virus interactions;

16 response to infection; specialization; virus evolution

17

18   **Abstract**

19   It is assumed that host genetic variability for susceptibility to infection will necessarily

20   condition the evolution of viruses, either by driving them to diversification into strains

21   that track the different host defense alleles (*e.g*., antigenic diversity), or by canalization

22   to infect only the most susceptible genotypes. Associated to these processes, virulence

23   may or may not increase. To tackle these questions, we performed evolution experiments

24   with turnip mosaic virus (TuMV) in *Arabidopsis thaliana* genotypes that differ in

25   mutations in genes involved in resistance pathways and in genes whose products are

26   essential for potyviruses infection. Plant genotypes classified into five groups according

27   to their degree of resistance and intensity of symptoms. We found that evolution

28   proceeded faster in the most resistant hosts than in the most permissive ones, as expected

29   for adaptation to a harsh environment. The multifunctional viral protein VPg turned out

30   to be the target of selection in most host genotypes. When all evolved TuMV lineages

31   were tested for fitness in all plant genotypes used in the experiments, we found that the

32   infection matrix was significantly nested, suggesting the evolution of generalist viruses

33   selected by the most restrictive mutant genotypes. At the other side, a modular pattern,

34   driven by convergent evolution of lineages evolved in the same host genotype, was also

35   observed.

36

The spectrum of disease severity can be attributed to heterogeneity in virus' virulence or in host's factors; the two are not necessarily independent explanations and they must actually complement and/or interact each other. A problem faced by viruses is that host populations consist of individuals that had different degrees of susceptibility to infection (Schmid-Hempel & Koella 1994; Pfenning 2001). 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 & Koella 1994; Altizer 2006; Hughes & Boomsma 2006; Brown & Tellier 2011; Anttila *et al*. 2015; Parrat *et al*. 2016; González *et al*. 2019). The interaction between host and parasite genotypes has been explained in the light of two different theoretical models that represent the two extremes in 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.

Viral infection of plants is a complex system in which the virus parasitizes the host and utilizes all its cellular resources to replicate and systemically spread. In response, plants have evolved intricated signaling mechanisms (Zhou & Zhang 2020) that limit the spread of the virus, resulting in resistance. A variety of factors contribute to plant

62 resistance to viral infections. Broadly speaking, these factors can be classified into *basal*

63 if they are pre-existing and limit within-cell propagation and cell-to-cell spread, and

64 *inducible* if they are only activated upon infection and inhibit systemic virus replication

65 and movement. Basal mechanisms include susceptibility (*S*) genes that involve alleles of

66 cellular proteins that do not interact properly with viral factors, specially translation

67 initiation factors, required by the virus for successful exploitation of the cell's protein

68 synthesis machinery, heat shock proteins, that assist them in the formation of multiple

69 multiprotein complexes, or DNA binding phosphatases (Truniger & Aranda 2009; Carr

70 *et al.* 2010; Mäkinen 2019). In contrast, inducible mechanisms include most of resistance

71 (*R*) genes whose expression result in a broad-scale change in plant physiology via diverse

72 signal transduction pathways, particularly those regulated by the hormones salicylic acid

73 (SA), jasmonic acid (JA) and ethylene (ET) (Soosaar *et al.* 2005; Carr *et al.* 2010). These

74 changes include local cell apoptosis (*e.g.*, hypersensitive responses - HR; Loebenstein

75 2009), the upregulation of nonspecific responses against many different types of

76 pathogens throughout the entire plant (systemic acquired resistance – SAR- and induced

77 systemic resistance – ISR) (Kachroo *et al.* 2006; Carr *et al.* 2010), and the activation of

78 the RNA-silencing-based resistance, which seems to play a role both in basal and

79 inducible mechanisms (Voinnet 2001; Carr *et al.* 2010). Early host responses following

80 virus detection include changes in ion fluxes (mainly Ca^{++}), activation of signaling

81 pathways, major alterations of transcriptomic profiles, generation of reactive oxygen

82 species (ROS) and production of nitric oxide (NO) (Soosaar *et al.* 2005). These

83 immediate changes are followed by HR and the recruitment of SA and JA/ET signaling

84 pathways. The SA-mediated defense signaling pathway results in SAR, while the JA/ET-

85 mediated defense signaling pathway results in ISR, the latter being specifically involved

86 in symbiotic interactions between plants and beneficial microbes. Indeed, it appears that

4

87 ISR is not effective against infection by most viruses (Ton *et al.* 2002; Loebenstein 2009;

88 Pieterse *et al.* 2009). Interestingly, both SAR and ISR pathways converge into two master

89 regulators, the *ENHANCED DISEASE SUSCEPTIBILITY 1* (*EDS1*) and the

90 *PHYTOALEXIN DEFICIENT 4* (*PAD4*) genes; *EDS1* and *PAD4* repress ISR and promote

91 SAR. Although the SA and JA/ET pathways have been classically viewed as mutually

92 antagonistic, several studies have revealed positive and negative crosstalk between them

93 (van Wees *et al.* 2000; Pieterse *et al.* 2012) as well as between them and the RNA-

94 silencing pathway (Soosaar *et al.* 2005; Carr *et al.* 2010; Yang *et al.* 2020). This

95 crosstalk, which has the network topological structure of an incoherent feed-forward loop,

96 creates robustness and tunability in the plant immune network (Mine *et al.* 2017).

97 A particularly well study example of *S* genes in the context of genetic variability

98 for resistance and virus adaptation in *Arabidopsis thaliana* (L.) Heynh are the

99 *RESTRICTION TO TEV MOVEMENT* (*RTM*) genes, composed by at least five different

100 loci (Sofer *et al.* 2017) that encode for a variety of functions needed by tobacco etch virus

101 (TEV; genus *Potyvirus*, family *Potyviridae*), and other potyviruses, to induce a systemic

102 infection of the plant. The mechanism of resistance in this case is well understood:

103 dominant alleles in all five loci result in resistance, whereas homozygous deletions in at

104 least one of the loci result in increased susceptibility to infection and different degrees of

105 systemic infection. By means of experimental evolution, it has been shown that TEV can

106 readily adapt to different plant ecotypes. In particular, serial passages in susceptible plant

107 genotypes *rtm1*/*rtm1* result in a viral isolate (TEV-*At*17) capable of infecting otherwise

108 fully resistant genotypes (Agudelo-Romero *et al.* 2008; Lalić *et al.* 2010). Additional

109 passages of TEV-*At*17 in plants with different *RTM* configurations resulted in increased

110 local adaptation and, more interestingly, in the evolution of a GFG-like interaction

111 mechanism: the most susceptible plant genotype selected for less pathogenic and highly

5

112   specialized viruses while the most resistant plant genotype selected for more virulent and

113   generalist TEV isolates (Hillung *et al*. 2015). Later on, these results were confirmed with

114   turnip mosaic virus (TuMV; genus *Potyvirus*, family *Potyviridae*) in the context of fully

115   monomorphic or maximally polymorphic experimental *A. thaliana* populations

116   (González *et al*. 2019).

117   However, these previous studies have barely paid any attention to the intricacies of

118   the underlying defense regulatory pathways of *A. thaliana*, the existence of multiple

119   defense responses (*e.g*., SAR, ISR, RNA-silencing, and the diversity of *S* genes) and the

120   potential crosstalk among them. Therefore, an open question is how this intricacy

121   determines the evolutionary fate of viruses. For instance, would two independent defense

122   pathways select for specialist viruses adapted to counteract each one thus conforming a

123   fitness tradeoff? Would a tradeoff in the investment of plant resources between pathways

124   (*e.g*., SAR *vs* ISR) result in a diversification of virus adaptive strategies? Would viruses

125   target crosstalk points across pathways thus evolving as generalists? Here we will use

126   the TuMV-*A. thaliana* experimental pathosystem to explore this sort of questions. After

127   exploring the variability in phenotypic responses to TuMV infection of a collection of *A.*

128   *thaliana* mutants in different *S* and *R* genes, we choose nine mutants which represent the

129   entire diversity in phenotypic responses, including both basal and inducible resistance

130   mechanisms. Then, we performed evolution experiments on each plant mutant genotype

131   and track the evolution of several disease-related traits. At the end of the evolution

132   experiment, we evaluated the effect of host genotypes in the rates of virus evolution and

133   the contribution of contingency, selection and stochasticity into the outcome of evolution.

134   Next, we explored whether the different TuMV lineages evolved as specialist or

135   generalists depending on the mutated defense mechanism of their local hosts. Finally,

136 we sought to identify the molecular changes experienced in the genome of the different

137 viral lineages and explored the possible adaptive value of a few convergent mutations.

138

139

140 **Results**

141 **Classification of *A. thaliana* genotypes according to their phenotypic response to**

142 **TuMV infection**

143 The 21 *A. thaliana* mutant genotypes used in the study are shown in Table 1, including

144 information about the affected signaling pathways or cellular processes as well as the

145 expected phenotype of infection relative to wild-type (WT) plants based on the

146 description of these mutants. Genotypes were classified according to their phenotypic

147 response to TuMV infection based in five different disease-related traits measured 18

148 days post-inoculation (dpi). Fig. 1 shows the neighbor-joining clustering of genotypes

149 according to their multi-trait phenotypic response to TuMV infection (data in Fig. S1);

150 we found five significant groups (hereafter named as G1 to G5). For all four members of

151 G1 the response was highly homogenous and consistent with an enhanced resistance

152 response to TuMV infection: no significant change in ΔDW whilst a significant and

153 consistent reduction in *SS*, *AUDPS*, *I*, and *VL* relative to infected WT plants. Three

154 members of G1 had enhanced SAR response (Table 1) and the fourth one (*i4g2*) is a well-

155 known *S* gene involved in plant resistance to potyvirus (Nicaise *et al*. 2007; Gallois *et al*.

156 2010). The only member of G2 was the strong apoptosis-inducer mutant *p58^{IPK}*, which

157 shows no significant changes in *SS*, *AUDPS*, *I*, and *VL* relative to infected WT plants but

158 a significant reduction in ΔDW, indeed, infected plants were heavier than the noninfected

159 controls, thus being this a case of increased tolerance to infection. The only member of

160 G3 was mutant *dbp2*, also described as specifically increasing resistance against another

161 potyvirus (Castelló *et al*. 2011), which shows a very interesting response to TuMV

162 infection: ΔDW was significantly increased while *SS*, *AUDPS* and *I* were significantly

163 reduced; no effect in *VL* was observed. Thus, *dbp2* represents an intermediate response

164 to TuMV infection between the highly resistant G1 and those genotypes in G4. G4

165 represents a sort of hotchpotch formed by genotypes that at most differ from WT plants

166 in two traits, but without a clear pattern. Remarkably, though, with the exception of *cpr5-*

167 *2*, *dcl2 dcl4*, and *dip2*, all show significant reductions in *VL*. Therefore, members from

168 G4 would be considered similar to WT in their phenotypic response to TuMV infection.

169 Finally, *jin1* was the only member of G5, which shows a significant increase in all

170 measured traits except *VL*, consistent with and enhanced susceptibility to TuMV

171 infection. This is a surprising finding since *jin1* has been described as inducing a

172 constitutive expression of the SA-dependent defenses that made plants extremely

173 resistant to the infection of biotrophic bacterial pathogens (Laurie-Berry *et al*. 2006). To

174 further characterize the progression of infection, we selected a subset of representatives

175 from groups G1 (*eds8-1*) and G4 (*cpr5-2*, *dcl2 dcl4* and *hsp90-1*) and evaluated daily *I*,

176 *SS* and *VL* (Fig. S2). The *cpr5-2* plants started showing symptoms faster, though only

177 80% of plants showed symptoms after 15 dpi. Only 20% of plants of genotype *eds8-1*

178 showed symptoms of infection, which were delayed on time. Infected *hsp90-1* and *dcl2*

179 *dcl4* plants were undistinguishable from the WT infected ones. Regarding *VL*, all

180 genotypes accumulated less virus than WT plants, with *cpr5-2* being the one that did so

181 the less (Fig. S2)

182 Representatives from the five groups were selected for the subsequent evolution

183 experiment, covering the entire spectrum of responses to TuMV infection and

184 mechanisms of resistance/susceptibility (shadowed rows in Table 1): *eds8-1* and *i4g2*

185 from G1 representing the hardest selection regimes (more resistant genotypes); *p58^{IPK}*

186   from G2; *dbp2* from G3; *cpr5-2*, *dcl2 dcl4*, *hsp90-1*, and *npr1-1* from G4; and *jin1* from

187   G5 representing the softest selective regime (less resistant genotype).

188

189   **Experimental evolution in the representative *A. thaliana* genotypes**

190   First, we sought to evaluate the dynamics of evolution for *AUDPS* (Fig. 2), *I* (Fig. 3) and

191   *VL* (Fig. 4). Data in these figures were fitted all together to the multivariate analysis of

192   covariance (MANCOVA) model described in Eq. 1 of the Methods section. Notice that

193   lineage *cpr5-2*/L2 showed a quite different evolution pattern from the other four lineages

194   evolved in *cpr5-2*. It never increased in either of the three traits and was lost after passage

195   five (Fig. 2, Fig. 3). Therefore, we ended up with 44 evolved lineages.

196   Table 2 shows the results of this MANCOVA. Despite a considerable amount of

197   noise in the time series, we found some significant results that can be summarized as

198   follows. First, a net significant and large ($\eta_P^2 = 0.735$) effect associated to the

199   evolutionary passages has been observed, indicating that the three phenotypic traits

200   related to TuMV infection evolved during the experiment. Second, plant genotypes had

201   a highly significant and large ($\eta_P^2 = 0.547$) effect on the phenotypic traits, suggesting that

202   TuMV evolutionary dynamics were strongly influenced by the local host genotype in

203   which it was being passaged. More interestingly, this effect associated to host plant

204   genotypes was strongly dependent on the passage number ($\eta_P^2 = 0.463$), thus suggesting

205   that the slope of the relationship between phenotypic traits and time was influenced by

206   the local host genotype. This effect will be further evaluated in the next section. Third,

207   independent lineages show a strong degree of evolutionary parallelism, as indicated by

208   nonsignificant differences among viral lineages evolved in the same local host genotype

209   (factors $L(G)$ and $t{\times}L(G)$). However, this last conclusion should be taken with certain

9

210   level of caution since the power of the two tests is relatively low ($1 - \beta < 0.800$) and

211   hence we might be wrongly failing to reject the null hypothesis (type II error).

212   A relevant question in evolutionary biology is the extent in which ancestral

213   differences determine the fate of evolution. In other words, the relative contribution of

214   adaptation, chance and historical contingency to the evolution or organismal fitness.

215   Following the logic exposed by Travisano *et al.* (1995), we can imagine two situations.

216   First, ancestral statistical differences among phenotypes could be preserved along

217   evolution despite a net increase in the mean trait values (due to selection) and differences

218   among replicated lineages (due to chance; *i.e.*, mutation and drift). In this situation, we

219   should expect a nonzero slope in a regression of the evolved phenotypic values against

220   the ancestral ones. The closer the slope to a value of one, the more importance of ancestral

221   differences. Second, if initial trait variation among ancestral genotypes was eliminated

222   from the evolved populations because the combined effect of adaptation and chance, we

223   should expect a regression slope smaller than one. The less effect of ancestral differences,

224   the flatter the slope, being zero in the extreme case of ancestral differences being

225   completely erased. To evaluate the role of the ancestral differences in *AUDPS*, *I* and *VL*

226   of TuMV across the nine different plant genotypes, we evaluated the magnitude of their

227   change at the end of the evolution experiment. Fig. 5 shows the plots of evolved *vs*

228   ancestral values. For illustrative purposes, the diagonals represent the null hypothesis of

229   absolute preservation of ancestral differences. In the case of *AUDPS* (Fig. 5A), a

230   significant regression exists between evolved and ancestral values ($R = 0.510$, $F_{1,42} =$

231   14.732, $P < 0.001$), with a slope of 0.296 ±0.077 (±1 SD). Since the slope is still

232   significantly different from zero, yet clearly flatter than the diagonal ($t_{42} = 9.112$, $P <$

233   0.001), we conclude that ancestral differences have been mostly removed by the

234   combined action of selection and chance yet not completely erased. In the case of *I* (Fig.

235 5B) the regression was not significant ($R = 0.087$, $F_{1,43} = 0.031$, $P = 0.568$), thus

236 suggesting that for this trait ancestral differences among *A. thaliana* genotypes have been

237 fully erased by adaptation and chance. Finally, in the case of *VL* (Fig. 5C), the regression

238 of evolved values on ancestral ones was significant ($R = 0.417$, $F_{1,43} = 8.822$, $P = 0.005$),

239 suggesting again that ancestral phenotypic differences have been largely removed by

240 selection and chance yet some still persist (regression slope -0.142 ± 0.048 significantly

241 different from one: $t_{42} = 23.903$, $P < 0.001$).

242

243 **Rates of phenotypic evolution for disease-related traits**

244 As mentioned in the previous section, we have observed a significant effect of *A. thaliana*

245 genotypes in the temporal evolution of the three phenotypic traits. To further explore this

246 effect, we estimated the rates of phenotypic evolution of *AUDPS* and *I* using the

247 ARIMA(1,0,0) model described in the Methods section (Eq. 2). *VL* was not included in

248 this analysis since the number of data points (three) in the time series was the same than

249 the number of parameters to be estimated, thus likely incurring in overfitting problems.

250 Fig. 6 shows the estimated rates of evolution, *β*, for host genotypes ordered from

251 the most resistant (G1) to the most permissive (G5) ones. Very interestingly, a significant

252 negative regression coefficient exists for both traits [*AUDPS*: $-1.460 \cdot 10^{-2} \pm 0.412 \cdot 10^{-2}$

253 ($R = 0.480$, $F_{1,42} = 12.561$, $P = 0.001$) and *I*: $-3.161 \cdot 10^{-3} \pm 0.909 \cdot 10^{-3}$ ($R = 0.473$, $F_{1,42} =$

254 12.105, $P = 0.001$)], suggesting that evolution always proceeded faster in the most

255 restrictive hosts and slower in the most permissive ones.

256

257 **Dissecting the role of specific defense mechanisms in TuMV evolution**

258 The rates of TuMV phenotypic evolution estimated in the previous section were further

259 analyzed in the context of the mutated defense signaling pathway encoded by *R* genes or

11

260   recessive resistances encoded by S genes. Rates of evolution β_{AUDPS} and β_I were fitted to

261   the MANOVA model shown in Eq. 3 of the Methods section. Fig. 7 graphically

262   summarizes the results of these analyses. First, we sought for differences in rates of

263   evolution between permissive (soft selection) and restrictive (hard selection) hosts (Fig.

264   7A). The multivariate analysis shows that lineages evolved in the more restrictive hosts

265   evolved significantly faster than those evolved in the more permissive hosts ($\Lambda = 0.327$,

266    $F_{2,34} = 34.923$, $P < 0.001$), with the effect being mostly driven by differences in the rate

267   of infectivity evolution β_I, as shown by the corresponding univariate analyses, being on

268   average 3.14 times faster in the most restrictive hosts. This observation confirms the

269   results shown in the previous section.

270   Next, we evaluated whether differences exist between the rates of phenotypic

271   evolution for viruses evolving in host genotypes carrying mutations affecting resistance

272   (R; $i.e.$, the SAR, ISR and RNA-silencing pathways) or susceptibility (S; $dbp2$, $dip2$,

273    $hsp90-1$, $i4g2$, and $p58^{IPK}$) genes (Fig. 7B). In this case, the multivariate analysis found

274   a small yet significant effect ($\Lambda = 0.790$, $F_{2,34} = 4.507$, $P = 0.018$), though none of the

275   two univariate analyses had enough power to detect it (indeed, changes went in opposite

276   directions: 1.02-fold larger for β_{AUDPS} in the S plants but 1.31 times larger for β_I in the R

277   plants).

278   Fig. 7C shows the comparison of rates of TuMV phenotypic evolution for lineages

279   evolving in plant genotypes mutant for the SA-dependent signaling defense pathways. In

280   this case we have three categories: mutations inducing a down-expression of the pathway,

281   mutations with no effect on this pathway and mutations resulting in overexpression.

282   Again, the MANOVA found highly significant differences among the three categories (Λ

283    $= 0.213$, $F_{4,68} = 19.861$, $P < 0.001$). The univariate analyses show that most of the

284   differences were due to a faster evolution of I, being 2.38 times faster in the host

12

285   genotypes over-expressing the SAR defenses than in those with a WT or even down-

286   expression of the SA signaling.

287   SA- and JA/ET-dependent pathways are considered as antagonistic defenses,

288   although more recent evidences point towards an active crosstalk between both (van Wees

289   *et al.* 2000; Pieterse *et al.* 2012; Mine *et al.* 2017). Under the antagonistic hypothesis,

290   we should expect the opposite trend than observed for the lineages evolved in plants over-

291   expressing SAR defenses (Fig. 7C). Fig. 7D shows the evolution rates data classified

292   according to whether host genotypes had mutations affecting the JA/ET defense (ISR)

293   signaling pathway. The MANOVA found a significant difference between the two

294   categories ($\Lambda = 0.433$, $F_{2,34} = 22.258$, $P < 0.001$), entirely driven by differences in β_I, as

295   confirmed by the univariate analyses. Indeed, as expected, the rates of evolution for

296   infectivity were 2-fold faster for TuMV lineages evolved in the ISR deficient hosts, in

297   agreement with the SAR/ISR antagonistic hypothesis.

298   Finally, we tested the effect of knocking down the host's RNA-silencing pathway

299   in the rates of TuMV evolution (Fig. 7E). In this case, the multivariate analysis also found

300   highly significant differences between rates of viral evolution in plants with a fully

301   functional RNA-silencing pathway and *dcl2 dcl4* plants ($\Lambda = 0.490$, $F_{2,34} = 17.729$, $P <$

302   0.001), with average β_I values being 21-fold larger in plants expressing a fully functional

303   RNA-silencing pathway than in *dcl2 dcl4* plants unable of producing siRNAs.

304   In conclusion from this section, we found that regardless the defense signaling

305   pathway, viruses evolving in permissive plants always evolve slower than their

306   counterparts evolving in more restrictive plant genotypes.

307

308   **Permissiveness of *A. thaliana* genotypes to infection drive the evolution of specialist**

309   **or generalist TuMV lineages**

310   To analyze the specificity of adaptation of each evolved TuMV lineage, we performed a

311   full cross-infection experiment in which all the 44 evolved lineages were inoculated into

312   ten plants of the nine *A. thaliana* genotypes used in the evolution experiments. The

313   presence of symptoms on each plant was evaluated daily for up to 12 dpi. These data

314   were analyzed using two different approaches, first an ANOVA-like that evaluates the

315   effect of host genotypes, viral lineages and their interactions (Schmid-Hempel, 2011);

316   and second, the inference and characterization of nestedness, modularity and

317   specialization of an infection network (Weitz *et al.* 2013). For the first approach we fitted

318   the presence/absence of symptoms data to the logistic regression model shown in Eq. 4

319   of the Methods section. The results are shown in Table 3. Despite its apparent

320   complexity, interpretation is straightforward. A *TG* factor describes the effect of

321   differences among host genotypes, irrespective of the infecting viral lineage; it thus

322   characterizes whether some host genotypes are more susceptible than other to infection.

323   As shown in Table 3, a *TG* has a highly significant effect, both by itself and in the

324   interaction with *t*. The magnitude of the net effect was very large ($\eta_P^2 = 0.569$), while its

325   interaction with *t* was rather small in magnitude ($\eta_P^2 = 0.010$). On average, the most

326   susceptible host genotypes were *jin1*, *dcl2 dcl4* and *cpr5-2* ($\bar{f} = 1.000$), whereas the most

327   resistant host genotypes were *p58^{IPK}* ($\bar{f} = 0.459 \pm 0.013$) and *eds8-1* ($\bar{f} = 0.233 \pm 0.013$).

328   Similarly, a highly significant *LG* effect exists by itself (of large magnitude $\eta_P^2 = 0.288$)

329   as well as in combination with *t* (though the magnitude of the interaction was rather small

330   in magnitude $\eta_P^2 = 0.005$), which means that the observation depends on the viral lineage

331   independently of the host genotype. On average, the lineages evolved in *eds8-1* ($\bar{f} =$

332   0.997 ± 23.710) and *dbp2* ($\bar{f} = 0.994 \pm 26.108$) were those inducing symptoms faster; by

333   contrast, lineages evolved in *i4g2* ($\bar{f} = 0.788 \pm 279.090$) showed symptoms much slower.

334   Finally, and most interesting, a highly significant interaction term *TG×LG* (as well as in

335   combination with t) indicates that the outcome depends on the particular combinations of

336   host genotypes and TuMV lineages. Indeed, the magnitude of the interaction effect was

337   large ($\eta_P^2 = 0.314$), though its dependence with dpi was not so relevant ($\eta_P^2 = 0.048$). For

338   example, in the extreme of the most susceptible genotypes, the case of *cpr5-2* perfectly

339   illustrates this case: lineages evolved in *cpr5-2*, *dbp2*, *dcl2 dcl4*, *eds8-1*, and *p58^IPK* all

340   show $\bar{f} = 1.000$, whereas lineages evolved in *i4g2* and *jin1* show much lower $\bar{f}$ values

341   (0.674 ±0.028 and 0.849 ±0.032, respectively). In the opposite extreme, *e.g.* in *eds8-1*

342   lineages evolved in *eds8-1* show $\bar{f} = 0.575$ ±0.050 while lineages evolved in *jin1* show $\bar{f}$

343   = 0.124 ±0.030; all other lineages showing intermediate values.

344   For the second analytical approach, we used the above infection dataset to estimate

345   the corresponding *AUDPS* values, thus creating an infection matrix of 44 rows [$L(LG)$]

346   by nine columns (TG). This matrix was transformed into a new binary matrix in which a

347   value of one meant that the *AUDPS* for a given TuMV lineage in a particular *TG* was

348   greater or equal than observed for its *LG* and zero otherwise. Fig. 8A shows the packed

349   infection matrix with rows and columns organized to better highlight is nestedness and

350   modularity (Weitz *et al*. 2013). The matrix could we used to build a binary infection

351   network (Fig. 8B) that provides the same information in a more graphical way. The

352   infection network was significantly nested ($T = 14.706$, $P < 0.001$), with the most

353   permissive genotypes *jin1* and *dcl2 dcl4* being susceptible to most viral lineages

354   (universally susceptible hosts), while the most resistant genotype *eds8-1* and *p58^IPK* were

355   only successfully infected by some lineages that were precisely evolved in these two

356   genotypes. Likewise, the most generalist viral lineages were *eds8-1*/L4, *p58^IPK*/L3 and

357   *p58^IPK*/L4 that infected with equal efficiency all nine host genotypes (universally virulent

358   viruses), while the most specialized viral lineages, able only to successfully infect their

359   local hosts, were all surviving lineages evolved in *cpr5-2*, four lineages evolved in *dcl2*

360 *dcl4* and four lineages evolved in *jin1* (Fig. 8). Therefore, we conclude that local host's

361 resistance to infection positively correlates with the host range of the evolved lineages:

362 more resistant host genotypes selected for very generalist viruses while less resistant host

363 genotypes selected for very specialized viruses, in agreement with a gene-for-gene

364 infection model (Weitz *et al.* 2013).

365 Finally, we computed Blüthgen *et al.* (2006) specialization indexes for the packed

366 matrix shown in Fig. 8. Firstly, we evaluated the standardized species-level measure of

367 partner diversity *d'*. *d'* ranged from zero for the most generalist TuMV lineages (*eds8-*

368 *1*/L4, *p58^{IPK}*/L3 and *p58^{IPK}*/L4) to one for the most specialist ones (the four surviving

369 *cpr5-2*-evolved lineages, *dbp2*/L2, *dcl2 dcl4*/L2 - L5 and *jin1*/L1 - L4) (last column in

370 Fig. 8A). Secondly, we computed the network-level specialization index H_2', obtaining a

371 value of $H_2' = 0$ (uncorrected $H_2 = 5.024$, $H_{min}' = 3.534$, $H_{max}' = 5.024$; $P < 0.001$), which

372 means that, overall, the binary infection network shows a degree of specialization, with

373 most lineages being able of infecting one or few host genotype, a result mainly driven by

374 those TuMV lineages evolved in the less resistant host genotypes.

375 Finally, we sought to explore if the infection network in Fig. 8B also shows

376 evidences of modularity. Indeed, from an ecological standpoint, a significant network-

377 level degree of specialization means that some TuMV lineages do not interact well with

378 some of the *A. thaliana* genotypes, thus creating the possibility of modularity. In this

379 context, a module will refer to an aggregated set of viral lineages characterized by more

380 interactions within the module than between modules (Newman 2006; Dormann &

381 Strauss 2014). To this end, we computed the *Q* modularity index, which ranges from

382 zero, when the community has no more links within modules than expected by chance,

383 to a maximum value of one. We found a small yet significant modularity ($Q = 0.234$, P

384 $= 0.047$) in the infection matrix (Fig. 8A). What are the implications of modularity in

16

385   TuMV pathogenesis? Modularity is expected for a MA infection model (Weitz *et al*.

386   2013). More similar host genotypes, *e.g.*, those with mutations affecting the same

387   signaling pathway, may be selecting for viral lineages with similar properties, thus being

388   able of infecting with equal efficiency the subset of plant genotypes. A less appealing

389   alternative hypothesis to explain the observed modularity may be parallel evolution of

390   lineages evolved into the same host genotype; the most representative cases of this

391   possibility being the lineages evolved in *cpr5-2*, *dcl2 dcl4* and *jin1*. To distinguish

392   between these two hypotheses, we computed a reduced matrix averaging the observed

393   *AUDPS* among TuMV lineages evolved in the same host genotype and transforming this

394   nine by nine infection matrix into a packed binary matrix as described above. If the

395   reduced matrix still shows significant modularity, the first hypothesis will hold, while if

396   modularity disappears by averaging lineages evolved in the same host genotype, then the

397   alternative hypothesis will be more parsimonious. The reduced matrix is still

398   significantly nested ($T = 13.064$, $P < 0.001$), but modularity was not significant anymore

399   ($Q = 0.286$, $P = 0.137$), thus supporting the alternative hypothesis that the observed

400   modularity was driven by convergent evolution of lineages evolved into the same host

401   genotype rather than by the overlap of selective pressures by genotypes carrying

402   mutations in the same defense mechanism.

403

404   **Genomic differences/similarities among TuMV lineages and VPg as a potential**

405   **target of selection**

406   Finally, we have explored the molecular changes experienced by the TuMV evolved

407   lineages. Full genomic consensus sequences were determined for all 44 evolved lineages.

408   A total of 114 mutational events have been observed, affecting 71 nucleotide position

409   (Supplementary Fig. S3A and Table S1). According to the type of nucleotide substitution

410   involved, 93 were transitions and 21 transversions; regarding their effect on the protein

411   sequence, 30 were synonymous and 84 nonsynonymous. Interestingly, some mutations,

412   including 13 nonsynonymous and four synonymous, have been observed multiple times

413   in independent lineages. As discussed in the Methods section, treating each lineage as an

414   observation and each host genotype as a subpopulation, the average nucleotide diversity

415   within host genotypes, referred only to the 71 polymorphic sites, is $\pi_S = 0.067 \pm 0.007$.

416   On the other hand, the nucleotide diversity for the entire sample is $\pi_T = 0.074 \pm 0.007$.

417   Consequently, the estimate of interhost genotypes nucleotide diversity is $\delta_{ST} = 0.007$

418   ± 0.003, and the estimated coefficient of nucleotide differentiation (Nei 1982), $N_{ST} = 0.090$

419   ± 0.033, a value significantly greater than zero ($z = 2.727$, $P = 0.003$). Thus, we conclude

420   that minor yet significant genetic differentiation has been generated among viral lineages

421   replicating in different host genotypes. To assess whether selection played a role in

422   genetic differentiation among *A. thaliana* genotypes, we performed a Tajima's *D* test

423   (Tajima 1989) and found that it was significantly negative ($D = -2.487$, $P = 0.006$).

424   Next, we sought to characterize the distribution of mutational events along the nine

425   non-overlapping cistrons (Fig. S3B). The frequency of mutations per cistron, relative to

426   the length of the corresponding cistron, was fitted to the logistic regression model shown

427   in Eq. 5 of the Methods. Highly significant differences exist ($\chi^2 = 127.545$, 9 d.f., $P <$

428   0.001), yet entirely due to the ~16-fold larger mutation frequency observed in the *VPg*

429   cistron relative to the rest of the genome (Fig. S3B). Notice that all mutations observed

430   in *VPg* are nonsynonymous and that all lineages except *eds8-1*/L3 carry at least one

431   mutation in this cistron (Fig. S3A and Table S1).

432   Convergent nonsynonymous mutations are, *a priori*, good candidates for adaptive

433   mutations. Mutation CP/V148I appears in two lineages evolved in *cpr5-2* (Fig. S3A and

434   Table S1), mutation CP/S70N appear in lineage *dcl2 dcl4*/L1 and three lineages evolved

18

435   in *jin1* (Fig. S3A and Table S1), and two different mutations affecting CP amino acid

436   112, CP/D112G (lineage *npr1-1*/L4) and CP/112A (lineage *i4g2*/L5), that result in a

437   similar replacement of side chains. In addition, seven nonsynonymous mutations in *VPg*

438   are shared by several lineages (Fig. S3A and Table S1). Out of these seven cases, three

439   seem particularly promising candidates. They all are located in a narrow region of VPg

440   (residues 113, 115 and 118) (Fig. S3A and Table S1). Firstly, mutations G6237A,

441   G6237C, A6238G, and U6239A all affect the same codon, resulting in amino acid

442   replacements D113N, D113H, D113G, and D113E, respectively. The expected

443   functional effect of these mutations was evaluated using the algorithms implemented in

444   SNAP2 (Hecht *et al*. 2015). D113E is predicted to be functionally neutral (score = −76,

445   accuracy = 87%), while D113G (32, 66%), D113H (43, 71%) and D113N (30, 66%) are

446   predicted to have a functional effect. Among these three amino acid replacements,

447   D113G, which shows an intermediate value of the functional effect score, is particularly

448   interesting, involving a very strong change in the side radical from a long negatively

449   charged one to a small nonpolar one. This replacement has been observed in lineages

450   *cpr5-2*/L3, *cpr5-2*/L4, *cpr5-2*/L5, *npr1-1*/L1, *npr1-1*/L5, and *hsp90-1*/L2. All three hosts

451   for these lineages belong to the phenotypic group G4 in Fig. 1. Secondly, mutations

452   A6243C, A6243G and A6244G also affect the same codon, resulting in amino acid

453   replacements N115H (−42, 72%), N115E (−66, 82%) and N115S (−30, 61%). All three

454   are predicted to be functionally neutral. Thirdly, mutations U6252 and G6253A affect

455   the same codon and result in amino acid replacements R118C (−10, 53%) and R118H

456   (56, 78%), respectively. The most frequent among these two replacements is R118H,

457   which represents a conservative change among large positively charged side radicals, yet

458   has a strong expected functional effect. R118H has been observed in four out of five

459   TuMV lineages evolved in the *A. thaliana* mutant *jin1* (phenotypic group G5 in Fig. 1).

460   D113G thus represents a case of convergent evolution non-specific of the local host

461   genotype (*i.e.*, a candidate for a generalist mutation) while R118H represents a mutation

462   highly specific of the local host genotype (*i.e.*, a specialist mutation).

463   To further characterize the possible adaptive value of mutations VPg/D113G and

464   VPg/R118H, we created by site-directed mutagenesis the two mutated versions of *VPg*

465   and cloned them into the ancestral p35STunos infectious clone (Chen *et al.* 2003).

466   Viruses were recovered from these clones and three disease-related phenotypic traits (*e.g.*,

467   *AUDPS*, *I* and *SS*) evaluated both in the WT plants as well as in their corresponding local

468   hosts (*i.e.*, *cpr5-2* for mutant VPg/D113G and *jin1* for mutant VPg/R118H). Fig. 9

469   summarizes the results of these experiments. Regarding VPg/D113G (Fig. 9A - C), the

470   results strongly depend on the combination of phenotypic trait and plant genotype: it

471   shows a significant negative effect in both *AUDPS* (-14.2%; $z = 1.687$, $P = 0.046$) and *I*

472   (-23.9%; $z = 6.344$, $P < 0.001$) but a positive one in *SS* (16.7%; $z = 2.597$, $P = 0.005$)

473   when evaluated in WT plants. However, it shows a largely negative yet not significant

474   effect in *cpr5-2* plants for *AUDPS* (-60.8%; $z = 1.055$, $P = 0.146$) and significant negative

475   effects both for *I* (-60.1%; $z = 10.614$, $P < 0.001$) and *SS* (-46.7%; $z = 3.228$, $P = 0.001$).

476   These negative effects on the local host genotype *cpr5-2* do not support a possible

477   beneficial effect of this mutation by itself. Regarding the possible adaptive effect of the

478   amino acid replacement VPg/R118H (Fig. 9D - F), the results are also dependent on the

479   combination of phenotypic trait and host plant genotype. When tested in WT plants, the

480   mutation has significant negative effects both for *AUDPS* (-11.4%; $z = 2.29$, $P = 0.013$)

481   and *I* (-9.0%; $z = 2.749$, $P = 0.003$) but a significantly positive effect for *SS* (43.6%; $z =$

482   7.881, $P < 0.001$). By contrast, in the local host *jin1* the effect was positive in all three

483   traits; significant for *AUDPS* (9.2%; $z = 1.683$, $P = 0.046$) and *SS* (68.2%; $z = 12.447$, P

484   < 0.001) but not for *I* (0.3%; $z = 1.342$, $P = 0.090$). These results support the conclusion

20

485   that the amino acid replacement VPg/R118H has an overall beneficial effect on the local

486   host genotype *jin1* in which it was selected.

487

488

489   Discussion

490   **Interaction between TuMV and the different defense pathways**

491   In this study we have explored the effect that mutations in host's different disease

492   signaling pathways (*R* genes) or in recessive resistance genes (*S*) have in the outcome of

493   virus evolution.   Our experimental pathosystem consisted in two well studied

494   components: the model plant *A. thaliana* and one virus with high prevalence in natural

495   populations on this host (Pagán *et al.* 2010), TuMV. Among the *R* genes, we have

496   included in our study genes involved in the SA, JA/ET and RNA-silencing pathways,

497   among the *S* genes, we have included heat shock proteins, transcription factors and

498   components of the translation machinery. In a preliminary set of experiments, we

499   stablished the similarities and differences in phenotypic responses of different plant

500   mutant genotypes to TuMV infection. Our results were, in some cases, at odds with the

501   expected phenotypes for *R* genes, being the results for *jin1* the most unexpected ones.

502   *JASMONATE INSENSITIVE 1* (*JIN1*) is a negative regulator of SA-mediated defense

503   responses, henceforth the mutant *jin1* has a constitutive expression of SAR (Laurie-Berry

504   *et al.* 2006). Surprisingly, it turns out to be the most sensitive genotype to TuMV

505   infection, showing enhanced symptoms of disease (phenotypic group G5). This may

506   reflect the fact that the function of *R* genes, especially those involved in SAR and ISR,

507   have been mostly defined in terms of plant interactions with biotrophic and necrotrophic

508   bacteria and fungi but rarely in response to virus infections. Mutant genotypes only

509   affecting components of the ISR pathway (*e.g.*, *coi1-4*) behaved as wild-type plants in

21

510　response to TuMV infection, thus confirming previous reports that ISR was inefficient

511　against viral infections (Ton et al. 2002; Loebenstein 2009; Pieterse *et al*. 2009).

512　However, mutant *eds8-1*, which avoids ISR but enhances SAR (Love *et al*. 2007) turned

513　out to be among the most resistant genotypes to infection (phenotypic group G1). In

514　agreement with our findings of a variable response to SA-signaling, Singh *et al*. (2004)

515　discussed examples of viruses in which SA-dependent responses to viral infection were

516　strongly dependent on the virus used in the experiments: while viruses such as alfalfa

517　mosaic virus, potato virus X or turnip vein clearing virus replication was inhibited by

518　treating plants with SA, cucumber mosaic virus was apparently unaffected by the

519　treatment and accumulated to normal levels.

520　　The RNA-silencing pathway is considered as the main plant defense against viral

521　pathogens (Voinnet 2001), with *DICER LIKE 2* (*DCL2*) and *DICER LIKE 4* (*DCL4*)

522　encoding for the two dicer enzymes responsible for generating the 22- and 21-nucleotides

523　long antiviral siRNAs. Therefore, the double mutant *dcl2 dcl4* was expected to show a

524　highly sensitive response to TuMV infection. However, infected plants were, overall,

525　hardly distinguishable from the WT plants in their response to infection (Fig. S2 and Fig.

526　S3). A possible explanation would be the strong suppressor activity of HC-Pro in the WT

527　plants that effectively counteracts the defense mechanism (Kaschau & Carrington 2001).

528　Indeed, it has been shown that the outcome of the interplay between plant RNA-silencing

529　response and potyvirus is mostly driven by the efficiency of the viral HC-Pro suppression

530　activity (Li *et al*. 2007; Torres-Barceló *et al*. 2008).

531　　The response to infection of genotypes carrying mutations in *S* genes was, in

532　general, consistent with the *a priori* expectation. Mutant *i4g2*, that is defective for the

533　eIF4(iso)E factor, shows a strong resistance to TuMV infection (Nicaise *et al*. 2007;

534　Charron *et al*. 2008), belonging to phenotypic group G1. Likewise, in agreement with

22

535   previous descriptions, mutant *dbp2* shows enhanced resistance to potyvirus infection

536   (Castelló *et al*. 2011), belonging to phenotypic group G3. Finally, mutant *hsp90-1* turns

537   out to show a response to TuMV infection equivalent to the WT plants, likely due to its

538   functional redundancy with other heat-shock proteins present in the cell that may be

539   coopted by the virus to assist in their expected functions (Verchot 2012).

540   An interesting finding was the effect on TuMV replication of knocking down the

541   expression of gen *P58^{IPK}*. In mammals P58^{IPK}, a tetratricopeptide repeat (TPR)-

542   containing protein, is recruited by viruses (*e.g*., influenza A virus) to inhibit interferon

543   activation and cell death mediated by the dsRNA-activated protein kinase (PKR), thus

544   favoring viral spread (Goodman *et al*. 2011). In our experiments, *p58^{IPK}* mutant plants

545   show enhanced tolerance to infection, showing weaker symptoms than WT plants despite

546   accumulating viruses at the same level (Fig. S1). In sharp contrast, Bilgin *et al*. (2003)

547   found that infection of *p58^{IPK}*-silenced *Nicotiana benthamiana* and *A. thaliana* plants

548   infected with tobacco mosaic virus lead to death. Given the limited information available

549   from other plant-virus pathosystems, additional experiments would be necessary to shed

550   light into the potential antiviral role in plants of this gene.

551   One side, yet important, conclusion of our study is that the generally accepted

552   model of plant defense signaling pathways, mainly based in experiments done with

553   bacteria and fungi, may not well describe the interaction between plants and viruses. By

554   contrast, recessive resistances based on *S* genes seem to better explain differences in

555   susceptibility to infection and viral accumulation, thus being more promising targets for

556   future development of resistant plants.

557

558   **Evolution of specialist and generalist viral strategies depend on the host genotype**

559   The GFG and MA models of host-virus interaction modes represent the two ends of a

560   continuum of possible outcomes (Agrawal & Lively 2002). The most relevant difference

561   between both models regards the expected genetic heterogeneity in both host and virus

562   populations. With a pure GFG interaction the susceptible host types are expected to

563   disappear and the resistant types will dominate the population. *Vice versa*, the most

564   virulent virus allele would drive to fixation at the cost of losing mild alleles. However,

565   constitutive activation of defenses is known to be costly for *A. thaliana* [*e.g.*, SA-related

566   defense responses pay a fitness cost in absence of pathogens (Traw *et al*. 2007)] and high

567   virulence usually comes with a cost in terms of pathogen's transmission (reviewed in

568   Acevedo *et al*. 2019). Hence, a pure GFG strategy seems unlikely to be achieved. By

569   contrast, with a pure MA interaction, negative frequency-dependent selection emerges,

570   such that rare *A. thaliana* resistance alleles have advantage and, as a result, a genetic

571   polymorphism shall be maintained (Schmid-Hempel 2011). Natural populations of *A.*

572   *thaliana* contain considerable amount of genetic variability for tolerance (Pagán *et al*.

573   2008) and for immunity-related genes (Todesco *et al*. 2010, van de Weyer *et al*. 2019),

574   thus suggesting that an arms race between pathogens and plants should be going on yet it

575   is unclear whether it may result in pure GFG or MA interactions or lie somewhere in

576   between. Evolution experiments with such different pathosystems as TuMV/*A. thaliana*

577   (González *et al*. 2019), *Octosporea bayeri*/*Daphnia magna* (Altermatt & Ebert 2008) and

578   *Serratia marcescens*/*Caenorhabditis elegans* (Gibson et al. 2020; White *et al*. 2020) have

579   produced congruent results in one aspect: parasites exposed to heterogeneous host

580   populations evolved significantly lower virulence than parasites exposed to homogeneous

581   host populations. However, a significant difference exists among pathosystems: while

582   viruses exposed to genetically heterogenous host populations evolved as no-cost

24

583   generalists, evolution of generalism in more complex parasites was constrained by a

584   fitness tradeoff, as expected for the jack-of-all trades hypothesis (Bedhomme *et al.* 2015).

585   Our results, as well as those by Hillung *et al.* (2014) and González *et al.* (2019)

586   have shown the evolution of significantly nested binary infection networks, a finding

587   compatible with the existence of a combination of specialist and generalist viruses and of

588   more permissive and resistant host genotypes. Indeed, these studies have also shown that

589   more permissive hosts selected for more specialized viruses while more resistant hosts

590   selected for more generalist viruses, here matching the predictions of the GFG model.

591   Our observation of small yet significant modularity in the infection network was easily

592   explained by convergent evolution of TuMV lineages evolved in the same host genotype.

593   However, it has been recently shown in a long-term survey of the prevalence of different

594   plant viruses in different hosts and habitats that nestedness and modularity in host-

595   pathogen infection networks is possible due to the spatially patched distribution of

596   habitats and temporal successions of plant species (Valverde *et al.* 2020): small spatial

597   scales create modularity that coexist with global nestedness. This pattern may change

598   spatially and temporally but remains stable over long evolutionary timescales.

599

Role of natural selection

601   We found evidences of significant genetic differentiation among TuMV lineages evolved

602   in different plant genotypes. To test whether these differences were driven by selection

603   we performed a Tajima's *D* test (Tajima 1989). The resulting *D* value was significative

604   and negative, which is compatible with the action of purifying selection, the presence of

605   slightly deleterious mutations segregating in the populations or fast population

606   expansions (Yang 2006). How to distinguish between these explanations? In

607   independent fast expanding populations, many new mutations may be generated and

25

608   rising in frequency in each population, thus being observed as singletons, mutations

609   present in only one of the many coexisting genomes in each evolving TuMV lineage.

610   Singletons inflate the number of segregating sites and thus cause $D < 0$. Indeed, this is

611   the case here: 54 out of the 71 observed variable sites are singletons, thus the observed

612   pattern of molecular diversity among lineages evolved in the same host genotype and in

613   different host genotypes is likely to be due to the fast expansions of viral populations.

614   However, we have found additional evidence supporting the action of positive

615   selection: the existence of a number of convergent nonsynonymous mutations arising in

616   independent lineages, some evolved in the same host genotype but some others rising in

617   different host genotypes (Table S1 and Fig. S3) yet without a clear association with the

618   particular signaling pathway or *S* gene being mutated (Table 1) nor the phenotypic cluster

619   they belong to (Fig. 1). Interestingly, most of these convergent mutations happened in

620   the *VPg* cistron, which turns out to be also the most variable one. VPg plays many

621   essential roles in genome transcription (it is linked to the 5'-end of the viral genome and

622   provides the hydroxyl group that primes the synthesis of the complementary strains by

623   the viral RdRp), translation (directly interacts with the eukaryotic initiation factors

624   eIF(iso)4E and eIF(iso)4G) and interacts with all other viral proteins (Bosque *et al*. 2014)

625   and some other cell proteins (Martínez *et al*. 2016). Indeed, in previous evolution

626   experiments with potyviruses, *VPg* has also been shown to be an important target of

627   selection. For example, Agudelo-Romero *et al*. (2008) found that a single amino acid

628   replacement in VPg was enough to largely increase TEV *I*, *VL* and *SS* in *A. thaliana*.

629   Similarly, Gallois *et al*. (2010) found that *A. thaliana* plants with knock-out mutations in

630   the *eIF(iso)4E*, *eIF(iso)4G1* and *eIF(iso)4G2* genes were resistant to TuMV infection.

631   Two mutations in the VPg (E116Q and N163Y) were enough to overcome this resistance

632   and return to the original infection phenotype, though yeast-two hybrid assays showed

633   that none of these mutations affected the binding of VPg with eIF(iso)4E. As a final

634   example, one of the most extensively used resistance genes in pepper commercial

635   cultivars against potato virus Y (PVY; genus *Potyvirus*, family *Potyviridae*) is the *pvr2*,

636   which has many different alleles (Nicaise *et al*. 2007; Charron *et al*. 2008). The *pvr2*

637   locus encodes for the eIF4E factor which, as mentioned above, physically interacts with

638   VPg.   Interestingly, all the resistance-breaking viral isolates found so far contain

639   mutations in the *VPg* cistron (Duprat *et al*. 2002; Moury *et al*. 2004; Ayme *et al*. 2006)

640   Interestingly, one of the two mutations identified by Gallois *et al*. (2010),

641   VPg/E116Q, affects the same protein domain that the two mutations we have identified

642   and characterized (D113G and R118H). The mechanism why these two mutations may

643   confer a selective advantage to TuMV lineages cannot be inferred from our studies. We

644   found VPg/D113G in several lineages evolved in different host genotypes, while

645   VPg/R118H was only found in the *jin1* lineages. Our assays failed to find a beneficial

646   effect of VPg/D113G but confirmed the host-specific beneficial effect of VPg/R118H in

647   *jin1* plants.

648   Unfortunately, due to economical constrains, we have not been able to exhaustively

649   testing all observed mutations and, hence, several others may still be potential candidates

650   for adaptive mutations.   More specifically, we have not tested possible epistatic

651   interactions among mutations fixed in the same genome. Three lineages carry more than

652   one nonsynonymous mutation in *VPg* (*dbp2*/L1 D113N and N115E, *eds8-1*/L1 N115H

653   and E116G, and *eds8-1*/L4 H33Y, D113N and K121E) (Table S1). In addition, 25

654   lineages carry additional nonsynonymous mutations in at least one other cistron besides

655   *VPg* (Table S1). Two lineages have up to three additional nonsynonymous mutations:

656   *cpr5-2*/L5 that has P3/T318M-T326S CI/V148I and VPg/D113G and *dcl2 dcl4*/L1 that

657   has P1/V54A CI/I378V VPg/R118C and CP/S70N. Therefore, plenty of opportunities

658   for epistatic effects both within and among cistrons exist that will be evaluated in future

659   experiments.

660

661

Methods

Plants, virus and growth conditions

664   A collection of 21 different *A. thaliana* mutants of the Col-0 accession were used for this

665   study (Table 1). In all experiments described below, plants were all maintained in a BSL-

666   2 climatic chamber under a photoperiod of 8 h light (LED tubes at PAR 90 - 100

667   μmol/m^2/s) at 25 ºC and 16 h dark at 20 ºC.

668   Prior to the inoculation experiments, we created a large stock of TuMV infectious

669   saps. Saps were obtained from TuMV-infected *N. benthamiana* Domin plants inoculated

670   with the infectious plasmid p35STunos that contains a cDNA of TuMV genome

671   (GeneBank accession AF530055.2) under the control of the cauliflower mosaic virus 35S

672   promoter and the *nos* terminator (Chen *et al.* 2003) as described elsewhere (González *et*

673   *al.* 2019; Corrêa *et al.* 2020). This TuMV sequence variant corresponds to the YC5

674   isolate from calla lily (*Zantesdeschia* sp) (Chen *et al.* 2003). After plants showed

675   symptoms of infection, they were pooled frozen with liquid N_2 and homogenized into a

676   fine powder using a Mixer Mill MM400 (Retsch GmbH, Haan, Germany). For

677   inoculations, the necessary amount of power was diluted in inoculation buffer (50 mM

678   phosphate buffer pH 7.0, 3% PEG6000, 10% Carborundum) and 5 μL of the inoculum

679   was gently rubbed into the plant leaves. Plants were all inoculated when reached growth

680   stage 3.5 in the Boyes' scale (Boyes 2001). This synchronization ensures that they all

681   were at the same phenological state when inoculated.

682

Phenotyping infected plants

The following five different traits were measured for each infected plant 18 dpi. (*i*) Change in dry weight (ΔDW) of the aerial part of infected plants, with a precision of 10 mg, relative to the corresponding noninfected controls (Fig. S1A). (*ii*) Severity of symptoms (*SS*) was evaluated in a semi-quantitative discrete scale (Fig 1B in Corrêa *et al.* 2020) ranging from zero for asymptomatic infections to four for plants showing a generalized necrosis and wilting (Fig. S1B). (*iii*) The area under the disease progress stairs (*AUDPS*) that summarizes the speed at which the disease severity progresses in a group of plants (Fig. S1C) and takes values in the range zero and the total number of plants included in the assay (Simko & Piepho 2012). (*iv*) Infectivity measured as the number of symptomatic plants out of the number of inoculated plants 18 dpi (*I*; Fig. S1D). And (*v*) viral load (*VL*) measured by absolute RT-qPCR as the number of viral genomes per ng of total RNA in the plant as described below (Fig. S1E).

The five traits were not all independent but showed some significant pairwise positive correlations: *AUDPS* with *I* ($r = 0.772$, 19 df, $P < 0.001$), *SS* ($r = 0.807$, 19 df, $P < 0.001$) and *VL* ($r = 0.472$, 19 df, $P = 0.031$), and *I* with *SS* ($r = 0.746$, 19 df, $P < 0.001$). Basically, ΔDW is orthogonal with the other four traits while the other four traits show some degree of association. Interesting is the case of *VL*, which is only (weakly) correlated to *AUDPS*.

Experimental evolution

Five TuMV lineages were evolved during 12 consecutive serial passages on each one of the nine selected mutant genotypes. To begin the evolution experiment, 10 21-days old *A. thaliana* plants per lineage, three leaves per plant, were inoculated as described above with the sap from *N. benthamiana*. Fourteen dpi, symptomatic *A. thaliana* plants were

708   harvested and used to prepare infectious saps as described above. These saps were 1/10

709   diluted in inoculation buffer and used to inoculate the next batch of plants.

710

711   **Total RNA extractions**

712   Tissue from each pool of infected plants per lineage, ecotype and serial passage was

713   collected, frozen with liquid N_2 and preserved at −80 ºC until it was homogenized into

714   fine powder using a Mixer Mill MM400. Next, an aliquot of approximately 100 mg of

715   grounded tissue per sample for which viral load was quantified, was used for total RNA

716   (RNAt) extraction performed with the Agilent Plant RNA isolation Mini kit (Agilent

717   Technologies, Santa Clara CA, USA). Aliquots of RNAt per each sample were separated

718   and their concentration adjusted at 50 ng/μL to estimate viral accumulation by RT-qPCR

719   (see below).

720

721   **Quantification of *VL***

722   *VL* of each plant sample per lineage, ecotype and selected passage was quantified by

723   absolute real-time quantitative RT-PCR (RT-qPCR) using standard curves and the

724   primers TuMV F117 forward (5'-CAATACGTGCGAGAGAAGCACAC-3') and F118

725   reverse (5'-TAACCCCTTAACGCCAAGTAAG-3') that amplify a 173 nucleotides

726   fragment from the *CP* cistron of TuMV genome, as previously described (Corrêa *et al*.

727   2020). Briefly, standard curves were constructed using ten serial dilutions of the TuMV

728   genome, that was synthesized by *in vitro* transcription as detailed previously (Cervera *et*

729   *al*. 2018), in RNAt extract obtained from healthy *A. thaliana* plants used as control in the

730   experiments. Amplification reactions were run in a 20 μL volume using the GoTaq 1-

731   Step RT-qPCR System (Promega, Madison WI, USA) and the recommended

732   manufacturer's instructions as described (Cervera *et al*. 2018) in an ABI StepOne Plus

733   Real-time PCR System (Applied Biosystems, Foster City CA, USA). The cycling

734   conditions consisted in: an RT phase of 5 min at 42 ºC and 10 min at 9 5ºC followed by

735   a PCR stage consisting in 40 cycles of 5 s at 95 ºC and 34 s at 60 ºC; and the final melt

736   curve profile analysis that consisted in 15 s at 95 ºC, 1 min at 60ºC and 15 s at 95 ºC.

737   Negative controls consisted in healthy RNAt plant extract (non infected control) and

738   water. Quantitative reactions were run as three technical replicates per sample and results

739   were analysed using the StepOne software 2.2.2 (Applied Biosystems).

740

741   **TuMV genome amplifications**

742   Evolved viral genomes of the passage 12 from each lineage and ecotype were amplified

743   by high-fidelity RT-PCR by using the AccuScript Hi-Fi (Agilent Technologies) reverse

744   transcriptase and Phusion DNA polymerase (Thermo Scientific, Waltham MA, USA)

745   following the manufacturer's instructions. Each complete TuMV genome was amplified

746   into three overlapping amplicons of 3114 (5' fragment R1), 3697 (central region R2) and

747   3287 nucleotides (3' fragment R3) using three primer sets. For RT reactions an aliquot

748   of the corresponding RNAt (150-300 ng) was mix with 0.25 μM of the 1R-P3 (5'-

749   CGAGTAGTATCTTATAGCACAGCGCTCCGACC-3'), 2R-NIa (5'-

750   TGTCTGGAATCGGTAGCAAATGTAGCTGAGTTGTG-3') or 3R-polyAR (5'-

751   TTTTTTTTTTTTTTTTTTTTGTCCCTTGCATCATATCAAATG-3') primer to

752   synthesize the R1, R2 or R3 cDNA fragment, respectively, that were denatured 5 min at

753   65 ºC and cooled on ice. Then it was added a mix containing AccuScript Hi-Fi 1× Buffer,

754   1 mM of dNTPs, 8 mM of DTT, 4U of Ribolock RNase inhibitor (Thermo Scientific) and

755   0.5 μL of AccuScript Hi-Fi (Agilent Technologies) in a 10 μL volume. RT reactions

756   consisted in a step of 90 min at 42 ºC to synthesize the cDNA followed by an incubation

757   of 5 min at 70 ºC to inactivate the enzyme. PCR reactions were performed in a 50 μL

31

758   volume containing a mix of 1× Phusion Buffer, 0.4 µM of dNTPs, 0.2 µM of each primer,

759   0.5-1 µL of DMSO, 2U of Phusion DNA polymerase (Thermo Scientific) and 1 µL of the

760   corresponding RT reaction. R1 fragment was amplified using the primer set 1F-5UTR

761   (5'-GCAAACGCAGACCTTTCGAAGCACTCAAGC-3') and 1R-P3 and the following

762   PCR conditions: an initial denaturation of 30 s at 98 ºC, 3 cycles of 10 s at 98 ºC, 20 s at

763   67 ºC and 2 min at 72 ºC, 3 cycles of 10 s a 98 ºC, 20 s at 65 ºC and 2 min at 72 ºC, and

764   32 cycles of 10 s at 98 ºC, 20 s at 63 ºC and 2 min at 72 ºC, followed by a final extension

765   step of 5 min at 72 ºC. Fragments R2 and R3 were amplified with primer set 2F-P3 (5'-

766   TGGGAGCTTGCGGATGGTGGATACACAATTC-3')   and   2R-NIa   or   3F-NIa   (5'-

767   CTCGTTATATGGAGTCGGTTTCGGACCACTCATAT-3')   and   3R-polyAR,

768   respectively and a PCR with the same denaturation and extension steps than the fragment

769   R1 but different amplification steps: an stage consisting in 15 cycles of 10 s at 98 ºC, 20

770   s 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

771   min at 72 ºC for fragment R2, while 15 cycles of 10 s at 98 ºC, 20 s at 67 ºC and 2 min at

772   72 ºC followed by 23 cycles of 10 s at 98 ºC, 20 s at 65 ºC and 2 min at 72 ºC for fragment

773   R3. PCR products were purified with the MSB Spin PCRapace Kit (Stratec Molecular,

774   Coronado CA, USA) and then Sanger-sequenced. Full-length consensus viral sequences

775   were obtained assembling the sequences of the three amplified products by using the

776   Genious R9.0.2 program.

777

778   **Construction of TuMV *VPg* mutants**

779   Mutants in the *VPg* cistron were synthesized by long inverse site-directed PCR

780   mutagenesis using the QuickChange II XL Kit (Stratagene, San Diego CA, USA)

781   following the manufacturer's instructions using p35STunos as template. Also, mutagenic

782   primers were designed following the manufacturer's recommendations. Primer set cpr5-

32

783   2F (5'-GGAGGATGAGTTGGGTCCAAATGAAATACGTGT-3') and cpr5-2R (5'-

784   ACACGTATTTCATTTGGACCCAACTCATCCTCC-3') was used to generate the

785   A6238G (D113G) mutant while primer set jin1F (5'-

786   GGATCCAAATGAAATACATGTGAATAAGACAATTC-3') and jin1R (5'-

787   GAATTGTCTTATTCACATGTATTTCATTTGGATCC-3') to obtain the G6253A

788   (R118H) mutant. PCR protocol consisted in: a denaturation step of 2 min at 95 ºC

789   followed by 20 cycles of 20 s at 95 ºC, 10 s at 60 ºC and 7 min at 68 ºC and a final

790   extension of 5 min at 68 ºC. After *Dpn*I digestion and transformation of electrocompetent

791   *Escherichia coli* DH5α, the presence of the desired mutations in the infectious clone and

792   the absence of undesired nucleotide changes was confirmed by sequencing.

793

794   **Bioassays of the TuMV *VPg* mutants**

795   The infectivity, viability and pathogenicity of both VPg TuMV mutants was confirmed

796   by performing three independent bioassays. In each experiment, three batches of 24 *A.*

797   *thaliana* plants (three-weeks old) of genotypes wild-type, *cpr5-2* and *jin1* were inoculated

798   with the wild-type TuMV plasmid clone used as reference and with plasmids of each of

799   the TuMV VPg mutants, while another two batches of 24 plants from *cpr5-2* or *jin1*

800   genotypes were inoculated with the WT TuMV plasmid and the corresponding plasmid

801   of the TuMV VPg mutant, the VPg/D113G mutant in the case of the *cpr5-2* ecotype and

802   that of VPg/R118H for the *jin1* genotype. *A. thaliana* were inoculated by abrasion of

803   three leaves applying equal amounts of the plasmid inoculum, a total of approximately 7

804   μg per plant distributed in 3 μL per leaf. Plasmids were previously purified by Midiprep

805   using the NucleoBond Xtra Midi Kit (Macherey-Nagel, Düren, Germany) and

806   resuspended in distilled water. Number of infected plants and symptom intensity were

807   collected every day until 15 dpi.

808

Statistical analyses

810 All statistical analyses described hereafter were performed with SPSS version 26 software

811 (IBM, Armonk, NY), unless otherwise indicated.

812 *AUDPS* (Fig. 2), *I* (Fig. 3) and *VL* (Fig. 4) data were fitted all together to a

813 multivariate analysis of covariance (MANCOVA) model in which plant genotype (*G*)

814 was the main factor, the independent evolution lineages (*L*) were nested within *G* and

815 passage (*t*) was introduced in the model as a covariable. The full model equation thus

816 reads:

817 $$\mathbf{P}_{ijk}(t) \sim \boldsymbol{\mu} + t + G_i + L(G)_{ij} + (t \times G)_i + [t \times L(G)]_{ij} + \boldsymbol{\varepsilon}_{ijk}, \qquad \text{(Eq. 1)}$$

818 where $\mathbf{P}_{ijk}(t) = (AUDPS_{ijk}, I_{ijk}, VL_{ijk})^{\mathrm{T}}$ is the vector of phenotypic traits observed at time *t*,

819 for an individual infected plant *k* of evolutionary lineage *j* of genotype *i*, $\boldsymbol{\mu}$ represents the

820 vector of phenotypic grand mean values and $\boldsymbol{\varepsilon}_{ijk}$ stands for the vector of errors assumed

821 to be Gaussian distributed at every *t*. The significance of each factor, the covariable and

822 their interactions was evaluated using the Wilks' *Λ* method. The magnitude of the effects

823 was evaluated using the η_P^2 statistic (proportion of total variability in the traits vector

824 attributable to each factor in the model; conventionally, values of $\eta_P^2 \geq 0.15$ are considered

825 as large effects).

826 Rates of phenotypic evolution for *AUDPS* and *I* were estimated by fitting the time

827 series data to a first-order autoregressive integrated moving-average, ARIMA(1,0,0),

828 model (Elena and Sanjuán, 2005; González *et al*. 2019). The model equation fitted has

829 the form:

830 $$Y_t - \rho_1 Y_{t-1} = Y_0 + \beta_Y t + \varepsilon_t, \qquad \text{(Eq. 2)}$$

831 where Y_k represents the variable being analyzed at passage *k*, ρ_1 measures the degree of

832 self-similarity in the time-series data (correlation between values at passages *t* and *t* − 1),

833   ε_t represents the sampling error at passage t, and β_Y represents the linear dependency of

834   variable Y with passage number, that is, the rate of phenotypic evolution. The rates of

835   phenotypic evolution were further analyzed in the context of the mutated defense

836   signaling pathway encoded by R genes or recessive resistances encoded by S genes. Rates

837   of phenotypic evolution were fitted to a multivariate analysis of variance (MANOVA) in

838   which the plant genotype serving as host for the evolutionary lineages (G) was nested

839   within the factor being analyzed (*i.e.*, the type of selection –hard *vs* soft-, the mode of

840   resistance –R *vs* S-, and whether mutations affected the SA-dependent, the JA/ET-

841   dependent signaling or the RNA-silencing pathways; Table 1). The full model equation

842   now thus reads:

843    $\mathbf{B}_{ijk} \sim \boldsymbol{\mu} + X_{ij} + X(G)_{ij} + \boldsymbol{\varepsilon}_{ijk},$ (Eq. 3)

844   where $\mathbf{B}_{ijk} = (\beta_{AUDPS}, \beta_I)^T$ is the vector of rates of phenotypic evolution observed for the

845   k replicate lineage evolved in plant genotype j within the factor X_i, as above, $\boldsymbol{\mu}$ represents

846   the vector of grand mean values and $\boldsymbol{\varepsilon}_{ijk}$ the vector of Gaussian errors.

847

848   **Analysis of infection the network**

849   The first statistical approach consisted in fitting a logistic regression model to the

850   presence/absence of symptoms data using GLM techniques with a Binomial probability

851   distribution and a probit link function. The model equation reads as follows:

852    $\text{probit}[f_{ijkl}(t)] \sim \phi + t + TG_i + t \times TG_i + LG_j + t \times LG_j + L(LG)_{jk} + t \times L(LG)_{jk} +$

853    $(TG \times LG)_{ij} + t \times (TG \times LG)_{ij} + [TG \times L(LG)]_{ijk} + t \times [TG \times L(LG)]_{ijk} + \varepsilon_{ijkl},$

854   (Eq. 4)

855   where $f_{ijkl}(t)$ is the frequency of symptomatic plants t dpi in the test host genotype TG, for

856   the viral lineage L evolved in the local host genotype LG. TG and LG were considered

857   as orthogonal factors, L was nested within LG and dpi (t) was treated as a covariable. The

858   model includes all main factors, their corresponding nested and factorial interactions as

859   well as their interactions with the covariable. ϕ represents the grand mean of the probit-

860   transformed f values, and ε the error term assumed to be Binomial. The significance of

861   the different factors was evaluated using likelihood-ratio tests (LRT).

862   For the second statistical approach, the binary infection matrix was analyzed using

863   tools borrowed from the field of network biology to explore whether they show random

864   associations between viral lineages and host genotypes, one-to-one associations,

865   nestedness indicative of a GFG type of interaction, or modularity (Weitz *et al.* 2013). The

866   statistical properties of the infection matrix were evaluated using the R package

867   "bipartite" version 2.15 (Dormann *et al.* 2008) in R version 4.0.0 (R Core Team 2020)

868   under RStudio 1.2.1335. Four different summary statistics were evaluated: T nestedness

869   (Bascompte *et al.* 2003), Q modularity (Newman 2006) and the d' species-level (or

870   Kullback-Leibler divergence) and H'_2 network-level (or two-dimensional normalized

871   Shannon entropy) specialization indexes (Blüthgen *et al.* 2006). H'_2 ranges between zero

872   and one for extreme generalists and specialists, respectively. Statistical significance of

873   these statistics was evaluated using Bascompte *et al.* (2003) null model.

874

875   **The statistics of molecular evolution**

876   Treating each lineage as an observation and each host genotype as a subpopulation, we

877   evaluated the average nucleotide diversity within host genotypes, π_S, the nucleotide

878   diversity for the entire sample, π_T, the interhost genotypes nucleotide diversity, δ_{ST}, and

879   the estimate of the proportion of interhost genotypes nucleotide diversity, known as

880   coefficient of nucleotide differentiation (Nei 1982), $N_{ST} = \delta_{ST}/\pi_T$. Standard deviations of

881   estimates were inferred from 1000 bootstrap samples. All these computations were done

882   using MEGA X (Kumar *et al.* 2018) and the lowest-BIC nucleotide substitution model

883   Kimura 2-parameters (Kimura 1980). Tajima's *D* test of selection (Tajima 1989) and its

884   statistical significance were evaluated using DnaSP6 (Rozas *et al*. 2017).

885   The frequency of mutations (*m*) per cistron (*C*), relative to the length of the

886   corresponding cistron, was fitted to the following logistic regression model using GLM

887   techniques with a Binomial probability distribution and a probit link function:

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

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

890   main ORF.

891   The functional effect of mutations found in the *VPg* cistron were evaluated *in silico*

892   using the SNAP2 algorithm (Hecht *et al*. 2015). The algorithm provides a functionality

893   score and an accuracy index for each possible mutation affecting a coding sequence.

894   Negative score values mean functional neutrality while positive values should be taken

895   as indications of functional effects; the larger the values, the stronger the effect.

896

References

898   Acevedo MA, Dillemuth FP, Flick AJ, Faldyn MJ, Elderd BD. 2019. Virulence-driven

899   trade-offs in disease transmission: a meta-analysis. *Evolution* **73**: 636-647.

900   Agrawal A, Lively CM. 2002. Infection genetics: gene-for-gene versus matching-alleles

901   models and all points in between. *Evol. Ecol. Res*. **4**: 79-90.

902   Agudelo-Romero P, Carbonell P, Pérez-Amador MA, Elena SF. 2008. Virus adaptation

903   by manipulation of host's gene experession. *PLoS ONE* **3**: e2397.

904   Altermatt F, Ebert D. 2008. Genetic diversity of Daphnia magna populations enhances

905   resistance to parasites. *Ecol. Lett*. **11**: 918-928.

906   Altizer S, Kaitala V, Laakso J, Ruokolainen L. 2006. Seasonality and the dynamics of

907   infectious diseases. *Ecol. Lett*. **9**: 467-484.

908   Anttila J, 2015. Environmental variation generates environmental opportunist pathogen

909   outbreaks. *PLoS ONE* **10**: e0145511.

910   Ayme V, Souche S, Caranta C, Jacquemond M, Chadoeuf J, Palloix A, Moury B (2006)

911   Different mutations in the genome-linked protein VPg of *Potato virus Y* confer

912   virulence on the *pvr2³* resistance in pepper. *Mol. Plant Microbe Interact.* **19**: 557-

913   563.

914   Bascompte J, Jordano P, Melián CJ, Olesen JM. 2003. The nested assembly of plant-

915   animal mutualistic networks. *Proc. Natl. Acad. Sci. USA* **100**: 9389-9387.

916   Bedhomme S, Hillung J, Elena SF. 2015. Emerging viruses: why are not jacks of all

917   trades? *Curr Opin Virol* **10**: 1-6.

918   Bilgin DD, Liu Y, Schiff M, Dinesh-Kumar SP. 2003. *P58^{IPK}*, a plant ortholog of double-

919   stranded RNA-dependent protein kinase PKR inhibitor, functions in viral

920   pathogenesis. *Dev. Cell* **4**: 651-661.

921   Blüthgen N, Menzel F, Blüthgen N. 2006. Measuring specialization in species interaction

922   networks. *BMC Ecol.* **6**: 9.

923   Bosque G, Folch-Fortuny A, Picó J, Ferrer A, Elena SF. 2014. Topology analysis and

924   visualization of *Potyvirus* protein-protein interaction network. *BMC Syst. Biol.* **8**:

925   129.

926   Bouché N, Lauressergues D, Gasciolli V, Vaucheret J. 2006. An antagonistic function for

927   *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral

928   siRNAs. *EMBO J.* **25**: 3347-3356.

929   Boyes DC. 2001. Growth stage-based phenotypic analysis of *Arabidopsis*: a model for

930   high throughput functional genomics in plants. *Plant Cell* **13**: 1499-1510.

931   Brown JKM, Tellier A. 2011. Plant-parasite coevolution: bridging the gap between

932   genetics and ecology. *Annu. Rev. Phytopathol.* **49**: 345-367.

933   Cao H, Bowling SA, Gordon AS, Dong X. 1994. Characterization of an *Arabidopsis*

934   mutant that is nonresponsive to inducers of systemic acquired resistance. *Plant Cell*

935   **6**: 1583-1592.

936   Carr JP, Lewsey MG, Palukaitis P. 2010. Signaling in induced resistance. *Adv. Virus Res*.

937   **76**: 57-121.

938   Castelló MJ, Carrasco JL, Navarrete-Gómez M, Daniel J, Granot D, Vera P. 2011. A plant

939   small polypeptide is a novel component of DNA-binding protein phosphatase 1-

940   mediated resistance to *Plum pox virus* in *Arabidopsis*. *Plant Physiol*. **157**: 2206-

941   2215.

942   Cervera H, Ambrós S, Bernet GP, Rodrigo G, Elena SF. 2018. Viral fitness correlates

943   with the magnitude and direction of the perturbation induced in the host's

944   transcriptome: the tobacco etch potyvirus-tobacco case study. *Mol. Biol. Evol*. **35**:

945   1599-1615.

946   Charron C, Nicolaï, M, Gallois JL, Robaglia C, Moury B, Palloix A, Caranta, C. 2008.

947   Natural variation and functional analyses provide evidence for coevolution between

948   plant eIF4E and potyviral VPg. *Plant J*. **54**: 56-68.

949   Chen CC, Chao CH, Chen CC, Yeh SD, Tsai HT, Chang CA. 2003. Identification of

950   *Turnip mosaic virus* isolates causing yellow stripe and spot on calla lily. *Plant Dis*.

951   **87**: 901-905.

952   Corrêa RL, Sanz-Carbonell A, Kogej Z, Müller SY, Ambrós S, López-Mogollón S,

953   Gómez G, Baulcombe DC, Elena SF. 2020. Viral fitness determines the magnitude

954   of transcriptomic and epigenomic reprogramming of defense responses in plants.

955   *Mol. Biol. Evol*. **37**: 1866-1881.

956　Cui H, Qiu J. Zhou Y, Bhandari DD, Zhao C, Bautor J, Parker JE. 2018. Antagonism of
957　　　transcription factor MYC2 by EDS1/PAD4 complexes bolsters salicylic acid defense
958　　　in *Arabidopsis* effector-triggered immunity. *Mol. Plant* **11**: 1053-1066.

959　Dormann CF, Gruber B, Fruend J. 2008. Introducing the bipartite package: analyzing
960　　　ecological networks. *R News* **8**: 8-11.

961　Dormann CF, Strauss R. 2014. A method for detecting modules in quantitative bipartite
962　　　networks. *Meth. Ecol. Evol*. **5**: 9-98.

963　Duprat A, Caranta C, Revers F, Menand B, Browning KS, Robaglia C. 2002. The
964　　　*Arabidopsis* eukaryotic initiation factor (iso)4E is dispensable for plant growth but
965　　　required for susceptibility to potyviruses. *Plant J*. **32**: 927-934.

966　Elena SF, Sanjuán R. 2005. RNA viruses as complex adaptive systems. *BioSystems* **81**:
967　　　31-41.

968　Flor HH. 1956. The complementary genetic systems in flax and flax rush. *Adv. Genet*. **8**:
969　　　29-54.

970　Frank SA. 1993. Specificity versus detectable polymorphism in host-parasite genetics.
971　　　*Proc. R. Soc. B* **254**: 191-197.

972　Gallois JL, Charron C, Sánchez F, Pagny G, Houvenaghel MC, Moretti A, Ponz F, Revers
973　　　F, Caranta C, German-Retana, S. 2010. Single amino acid changes in the turnip
974　　　mosaic virus viral genome-linked protein (VPg) confer virulence towards
975　　　*Arabidopsis thaliana* mutants knocked out for eukaryotic initiation factors
976　　　eIF(iso)4E and eIF(iso)4G. *J. Gen. Virol*. **91**: 288-293.

977　Gibson AK, Baffoe-Bonnie H, Penley MJ, Lin J, Owens R, Khalid A, Morran LT. 2020.
978　　　The evolution of parasite host range in heterogeneous host populations. *J. Evol. Biol*.
979　　　33: 773-782.

980   González R, Butković A, Elena SF. 2019. Role of host genetic diversity for susceptibility-
981   to-infection in the evolution of virulence of a plant virus. *Virus Evol.* **5**: vez024.

982   Goodman AG, Tanner BCW, Chang ST, Esteban M, Katze MG. 2011. Virus infection
983   rapidly activates the P58[IPK] pathway, delaying peak kinase activation to enhance viral
984   replication. *Virology* **417**: 27-36.

985   Gupta V, Willits MG, Glazebrook J. 2000. *Arabidopsis thaliana EDS4* contributes to
986   salicylic acid (SA)-dependent expression of defense responses: evidence for
987   inhibition of jasmonic acid signaling by SA. *Mol. Plant Microbe Interact.* **13**: 503-
988   511.

989   Hecht M, Bromberg Y, Rost B. 2015. Better prediction of functional effects for sequence
990   variants. *BMC Genomics* **16**: S1.

991   Hillung J, Cuevas JM, Valverde S, Elena SF. 2014. Experimental evolution of an
992   emerging plant virus in host genotypes that differ in their susceptibility to infection.
993   *Evolution* **68**: 2467-2480.

994   Hughes WHO, Boomsma JJ. 2006. Does genetic diversity hinder parasite evolution in
995   social insect colonies? *J. Evol. Biol.* **19**: 132-143.

996   Kachroo P, Chandra-Shekara AC, Klessig DF. 2006. Plant signal transduction and
997   defense against viral pathogens. *Adv. Virus Res.* **66**: 161-191.

998   Kaschau K, Carrington JC. 2001. Long-distance movement and replication maintenance
999   functions correlate with silencing suppression activity of potyviral HC-Pro. *Virology*
1000   **285**: 71-81.

1001   Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions
1002   through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111-120.

1003   Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary
1004   genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**: 1547-1549.

1005   Lalić J, Agudelo-Romero P, Carrasco P, Elena SF. 2010. Adaptation of tobacco etch

1006   Potyvirus to a susceptible ecotype of *Arabidopsis thaliana* capacitates it for systemic

1007   infection of resistant ecotypes. *Phil. Trans. R. Soc. B* **65**: 1997-2008.

1008   Laurie-Berry N, Joardar V, Street IH, Kunkel BN. 2006. The *Arabidopsis thaliana*

1009   *JASMONATE INSENSITIVE 1* gene is required for suppression of salicylic acid-

1010   dependent defenses during infection by *Pseudomonas syringae*. *Mol. Plant Microbe*

1011   *Interact.* **19**: 789-800.

1012   Lin SS, Wu HW, Jan FJ, Hou RF, Yeh SD. 2007. Modifications of the helper component-

1013   protease of *Zucchini yellow mosaic virus* for generation of attenuated mutants for

1014   cross protection against severe infection. *Phytopathology* **97**: 287-296.

1015   Loebenstein G. 2009. Local lesions and induced resistance. *Adv. Virus Res.* **75**: 73-117.

1016   Love AJ, Laval V, Geri C, Laird J, Tomos AD, Hooks MA, Milner JJ. 2007. Components

1017   of *Arabidopsis* defense- and ethylene-signaling pathways regulate susceptibility to

1018   *Cauliflower mosaic virus* by restricting long-distance movement. *Mol. Plant Microbe*

1019   *Interact.* **20**: 659-670.

1020   Mäkinen K. 2019. Plant susceptibility genes as a source for potyvirus resistance. *Ann.*

1021   *Appl. Biol.* **173**: 122-129.

1022   Martínez F, Rodrigo G, Aragonés V, Ruiz M, Lodewijk I, Fernández U, Elena SF, Daròs

1023   JA. 2016. Interaction network of tobacco etch potyvirus NIa protein with the host

1024   proteome during infection. *BMC Genomics* **17**, 1

1025   Mine A, Nobori T, Salazar-Rodon MC, Winkelmüller TM, Anver S, Becker D, Tsuda K.

1026   2017. An incoherent feed-forward loop mediates robustness and tunability in a plant

1027   immune network. *EMBO Rep.* **18**: 464-476.

1028   Moury B, Morel C, Johansen E, Guilbaud L, Souche S, Ayme V, Caranta C, Palloix A,

1029   Jacquemond M. 2004. Mutations in Potato virus Y genome-linked protein determine

1030   virulence towards recessive resistances in *Capsicum annuum* and *Lycopersicon*

1031   *hirsutum*. *Mol. Plant Microbe Interact*. **17**: 322-329.

1032   Nawrath C, Heck, S, Parinthawong N, Métraux JP. 2002. *EDS5*, an essential component

1033   of salicylic acid-dependent signaling for disease resistance in *Arabidopsis*, is a

1034   member of the MATE transporter family. *Plant Cell* **14**: 275-286.

1035   Nawrath C, Métraux JP. 1999. Salicylic acid induction-deficient mutants of *Arabidopsis*

1036   express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen

1037   inoculation. *Plant Cell* **11**: 1393-1404.

1038   Nei M. 1982. Evolution of human races at the gene level. In *Human Genetics, part A: the*

1039   *Unfolding Genome* (ed. B. Bonné-Tamir). Alan R. Liss, New York.

1040   Newman MEJ. 2006. Modularity and community structure in networks. *Proc. Natl. Acad.*

1041   *Sci. USA* **103**: 8577-8582.

1042   Nicaise V, Gallois JL, Chafiai F, Allen LM, Schurdi-Levraud V, Browning KS,

1043   Candresse T, Caranta C, Le Gall O, German-Retana S. 2007. Coordinated and

1044   selective recruitment of eIF4E and eIF4G factors for potyvirus infection in

1045   *Arabidopsis thaliana*. *FEBS Lett*. **581**: 1041-1046.

1046   Pagán I, Alonso-Blanco C, García-Arenal F. 008. Host responses in life-history traits and

1047   tolerance to virus infection in *Arabidopsis thaliana*. *PLoS Pathog*. **4**: e1000124.

1048   Pagán I, Fraile A, Fernández-Fueyo E, Montes N, Alonso-Blanco C, García-Arenal F.

1049   2010. *Arabidopsis thaliana* as a model for the study of plant-virus co-evolution. *Phil.*

1050   *Trans. R. Soc. B* **365**: 1983-1995.

1051   Parrat SR, Numminen E, Laine AL. 2016. Infectious disease dynamics in heterogeneous

1052   landscapes. *Annu. Rev. Ecol. Evol. Syst*. **47**: 283-306.

1053   Pfenning K. 2001. Evolution of pathogen virulence: the role of variation in host

1054   phenotype. *Proc. R. Soc. B* **268**: 755-760.

1055   Pieterse CMJ, León-Reyes A, van der Ent S, van Wees SC. 2009. Networking by small-

1056   molecule hormones in plant immunity. *Nat. Chem. Biol*. **5**: 308-316.

1057   Pieterse CMJ, van der Does D, Zamioudis C, León-Reyes A, van Wees SCM. 2012.

1058   Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol*. **28**: 489-521.

1059   Rozas J, Ferrer-Mata A, Sánchez-del Barrio JC, Guirao-Rico S, Librado P, Ramos-Osins

1060   SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of

1061   large data sets. *Mol. Biol. Evol*. **34**: 3299-3302.

1062   Schmid-Hempel P. 2011. *Evolutionary Parasitology*. Oxford University Press, Oxford:

1063   UK, pp. 259-272.

1064   Schmid-Hempel P, Koella JC. 1994. Variability and its implications for host-parasite

1065   interactions. *Parasitol. Today* **10**: 98-102.

1066   Simko I, Piepho HP. 2012. The area under the disease progress stair: calculation,

1067   advantage, and application. *Phytopathology* **102**: 381-389.

1068   Singh DP, Moore CA, Gilliland A, Carr JP. 2004. Activation of multiple antiviral defence

1069   mechanisms by salicylic acid. *Mol. Plant Pathol*. **5**: 57-63.

1070   Sofer L, Cabanillas D, Gayral M, Téplier R, Pouzoulet J, Ducousso M, Dufin L, Bréhélin

1071   C, Ziegler-Graff V, Brault V, Revers Frédéric. 2017. Identification of host factors

1072   potentially involved in *RTM*-mediated resistance during potyvirus long distance

1073   movement. *Arch. Virol*. **162**: 1855-1865.

1074   Soosaar JLM, Burch-Smith TM, Dinesh-Kumar SP. 2005. Mechanisms of plant

1075   resistance to viruses. *Nat. Rev. Microbiol*. **3**: 789-799.

1076   Tajima F. 1989. Statistical methods for testing the neutral mutation hypothesis by DNA

1077   polymorphism. *Genetics* **123**: 585-595.

1078   Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, Nomura K, He SY, Howe

1079   GA, Browse J. 2007. JAZ repressor proteins are targets of the SCF(COI1) complex

1080   during jasmonate signalling. *Nature* **448**: 661-665.

1081   Todesco M, Balasubramanian S, Hu TT, Traw MB, Horton M, Epple P, Kuhns C,

1082   Sureshkumar S, Schwartz C, Lanz C, Laitinen RAE, Huang Y, Chory J, Lipka V,

1083   Borevitz JO, Dangl JL, Bergelson J, Nordborg M, Weigel D. 2010. Natural allelic

1084   variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* **465**:

1085   632-636.

1086   Ton J, van Pelt JA, van Loon LC, Pieterse CMJ. 2002. Differential effectiveness of

1087   salicylate-dependent and jasmonate/ethylene-dependent induced resistance in

1088   *Arabidopsis*. *Mol. Plant Microbe Interact*. **15**: 27-34.

1089   Torres-Barceló C, Martín S, Daròs JA, Elena SF. 2008. From hypo- to hypersuppression:

1090   effect of amino acid substitutions on the RNA-silencing suppressor activity of the

1091   tobacco etch potyvirus HC-Pro. *Genetics* **180**: 1039-1049.

1092   Travisano M, Mongold, JA, Bennett, AF, Lenski RE. 1995. Experimental test of the roles

1093   of adaptation, chance, and history in evolution. *Science* **267**: 87-90.

1094   Traw MB, Kniskern JM, Bergelson J. 2007. SAR increases fitness of *Arabidopsis*

1095   *thaliana* in the presence of natural bacterial pathogens. *Evolution* **61**: 2444-2449.

1096   Truniger V, Aranda MA. 2009. Recessive resistance to plant viruses. *Adv. Virus Res*. **75**:

1097   119-159.

1098   Valverde S, Vidiella B, Montáñez R, Fraile A, Sacristán S, García-Arenal F. 2020.

1099   Coexistence of nestedness and modularity in host-pathogen infection networks. *Nat.*

1100   *Ecol. Evol*. **4**: 568-577.

1101   Van de Weyer AL, Monteiro F, Furzer OJ, Nishimura MT, Cevik V, Witek K, Jones JDG,

1102   Dangl JL, Weigel D, Bemm F. 2019. A species-wide inventory of NLR genes and

1103   alleles in *Arabidopsis thaliana*. *Cell* **178**: 1260-1272.

1104   Van Wees SCM, de Swart EAM, van Pelt JA, van Loon, LC, Pieterse CMJ. 2000.

1105   Enhancement of induced disease resistance by simultaneous activation of salicylate-

1106   and jasmonate-dependent defense pathways in *Arabidopsis thaliana*. *Proc. Natl.*

1107   *Acad. Sci. USA* **97**: 8711-8716.

1108   Verchot J. 2012. Cellular chaperones and folding enzymes are vital contributors to

1109   membrane bound replication and movement complexes during plant RNA virus

1110   infection. *Front. Plant Sci*. **3**, 271.

1111   Voinnet O. 2001. RNA silencing as a plant immune system against viruses. *Trends Genet*.

1112   **17**: 449-459.

1113   Weitz JS, Poisot T, Meyer JR, Flores CO, Valverde S, Sullivan MB, Hochberg ME. 2013.

1114   Phage-bacteria infection networks. *Trends Microbiol*. **21**: 82-91.

1115   White PS, Choi A, Pandey R, Menezes A, Penley M, Gibson AK, de Roode J, Morran L.

1116   2020. Host heterogeneity mitigates virulence evolution. *Biol. Lett* **16**: 20190744.

1117   Yang Z. 2006. *Computational Molecular Evolution*. Oxford University Press. Oxford:

1118   UK, pp 262-264.

1119   Yang Z, Huang Y, Yang J, Yao S, Zhao K, Wang D, Qin Q, Bian Z, Li Y, Lan Y, Zhou

1120   T, Wang H, Liu C, Wang W, Qi Y, Xu Z, Li Y. 2020. Jasmonate signaling enhances

1121   RNA silencing and antiviral defense in rice. *Cell Host Microbe* **28**: 89-103.

1122   Zhou JM, Zhang Y. 2020. Plant immunity: danger perception and signaling. *Cell* **181**:

1123   978-989.

1124

1125   **Data availability**

1126 Raw data will be available upon request.

1127

Acknowledgments

1133

Author contributions

1135 S.F.E. conceived the study, analyzed the data and wrote the paper. R.N., S.A., F.M.,

1136 B.W., and J.L.C. performed all the experiments.

1137

Competing interests

1139 The authors declare no competing interests.
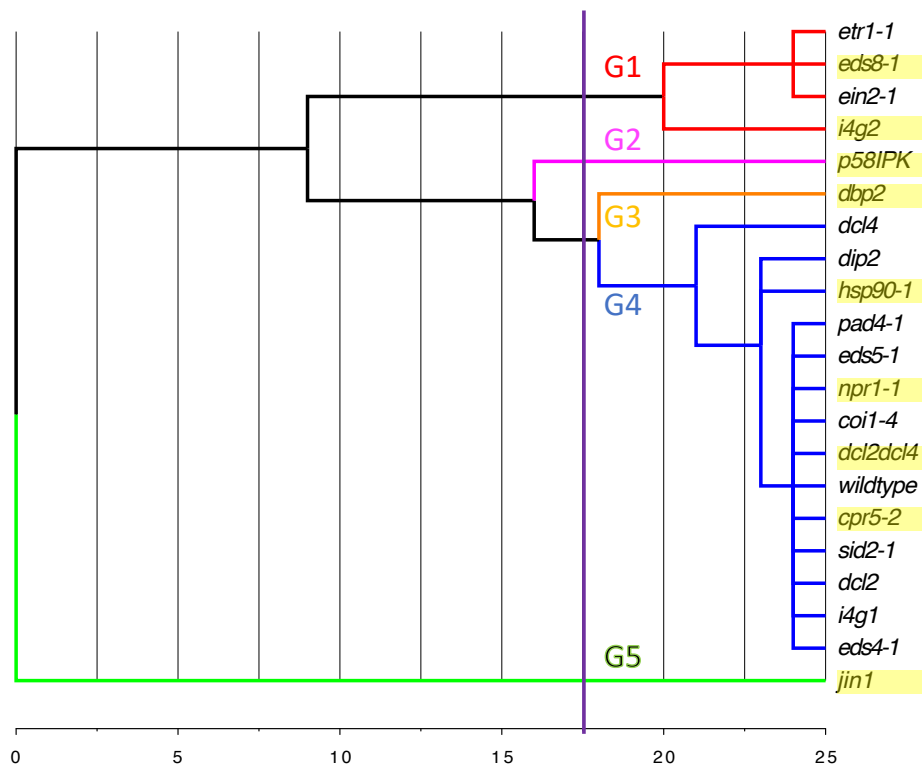
1140

1141 **Figures and Tables**

1142

Fig. 1. Neighbor-joining clustering of the 21 mutant genotypes of Arabidopsis according to their phenotypic similarity in response to TuMV infection. Genotypes selected as hosts for the evolution experiments are highlighted in yellow.
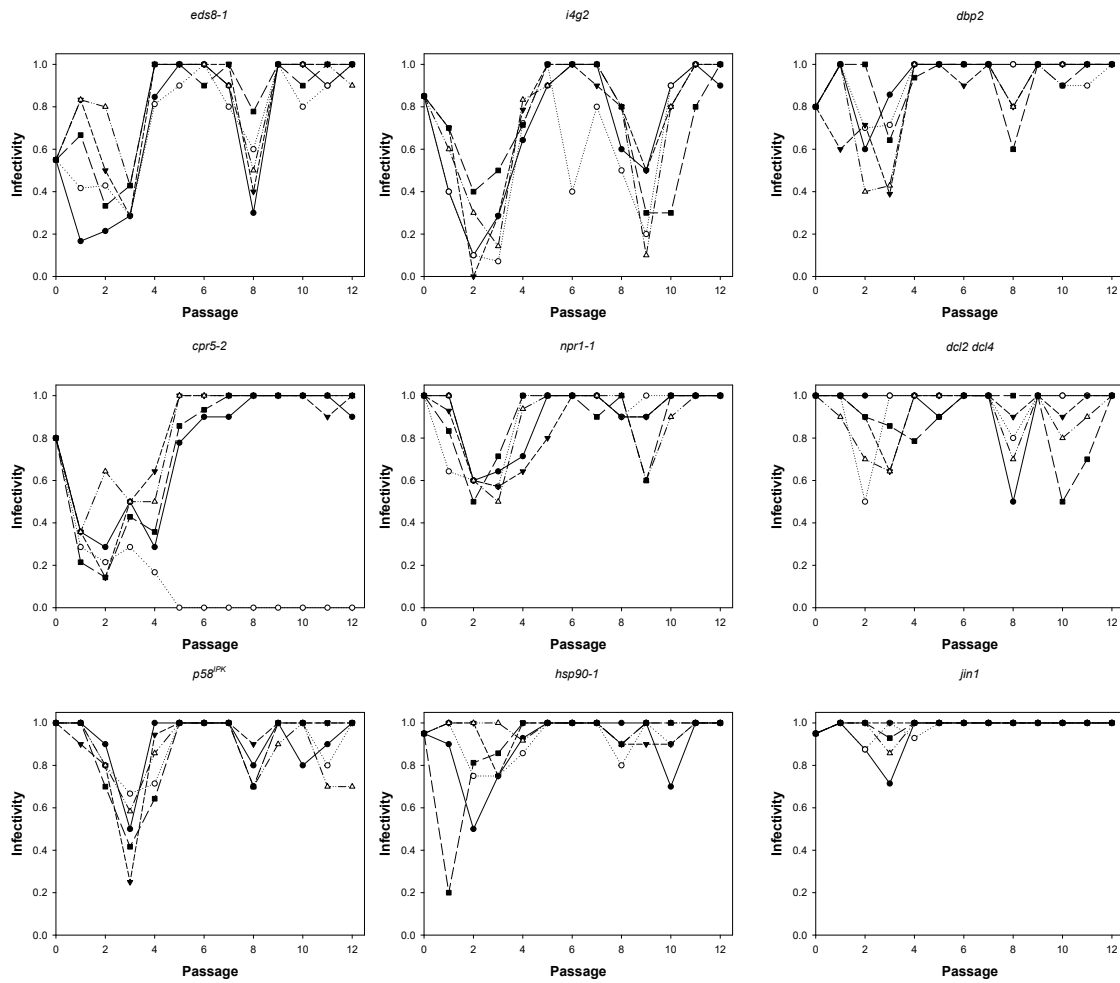
1143
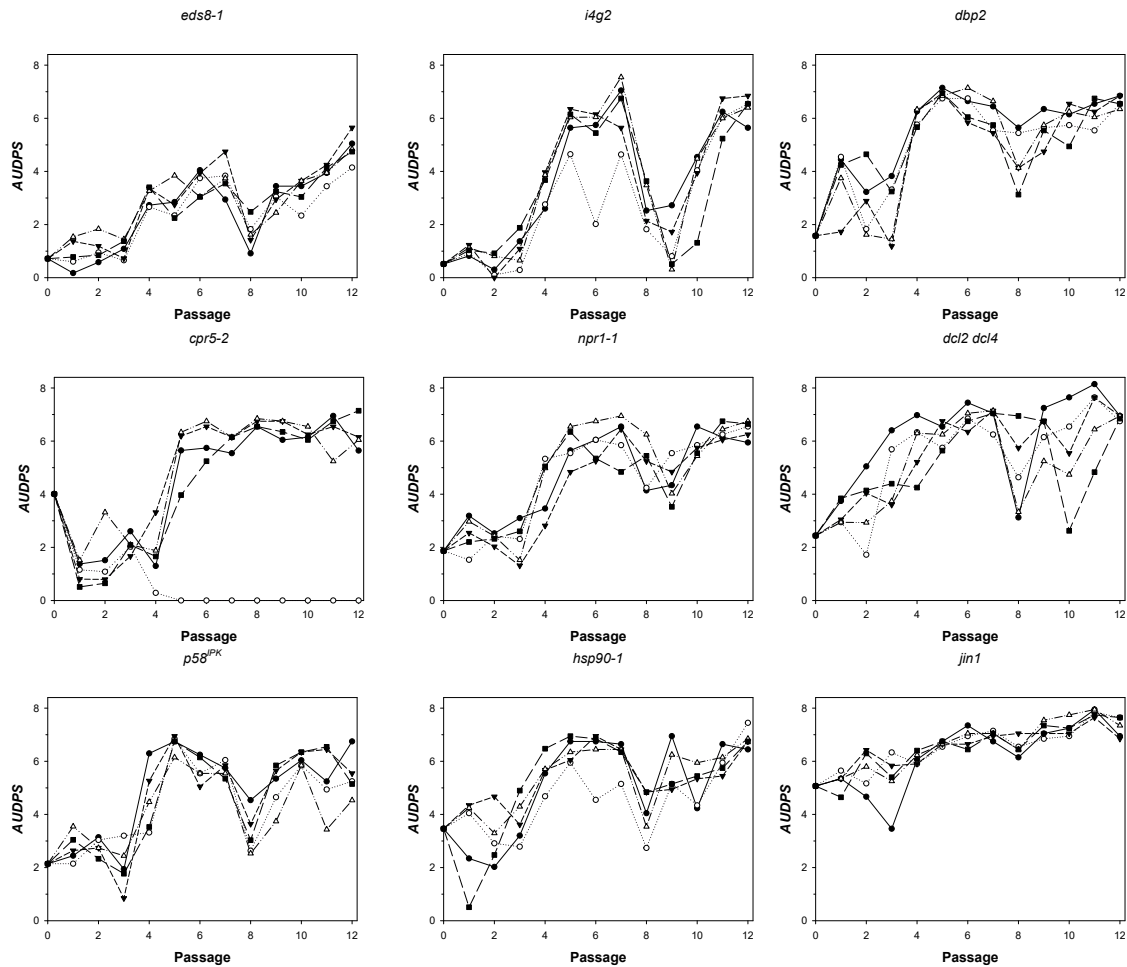
Fig. 2. Evolution of disease progression (*AUDPS*) along the serial passages of experimental evolution on each different host genotype. Different symbols and lines represent the independent evolutionary lineages. Panels are arranged from the most resistant genotype (*eds8-1*) to the most sensitive one (*jin1*) according to the groups defined in Fig. 1.
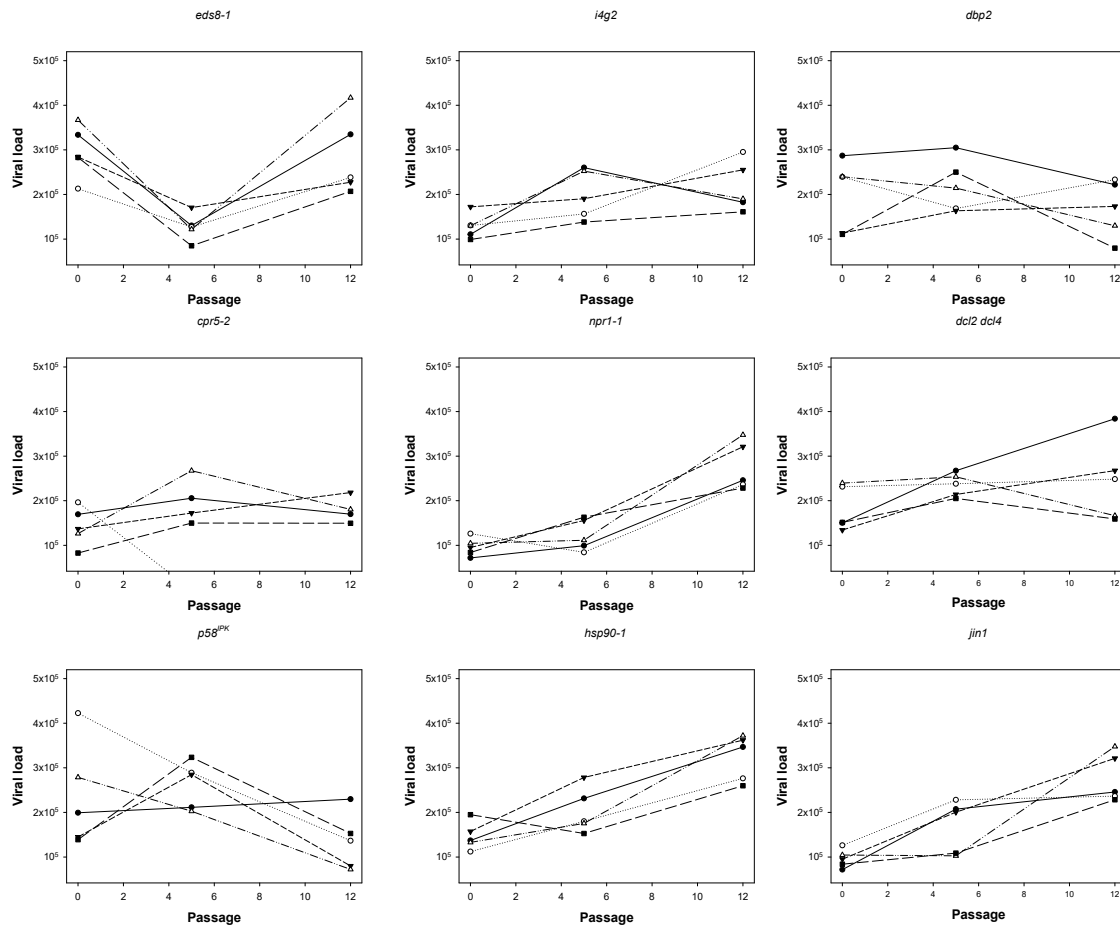
1144

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 Fig. 1.

1145

50

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 Fig. 1.
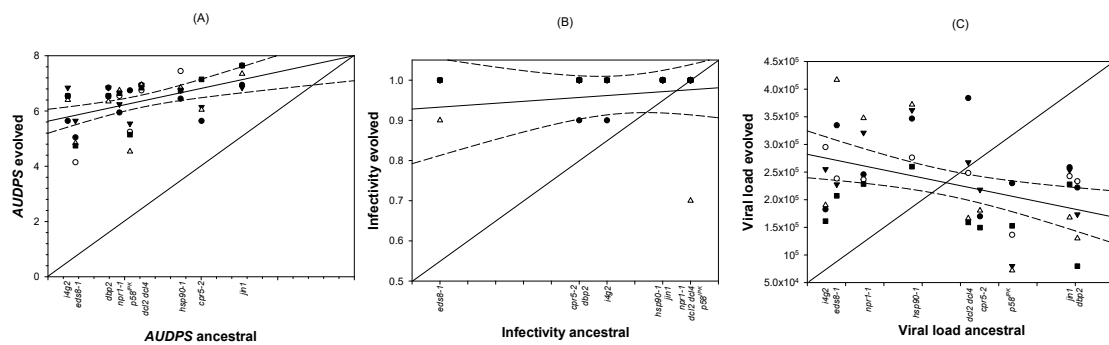
1146

Fig. 5. Test of the contribution of historical contingency to the observed pattern of adaptation of TuMV to the different Arabidopsis 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).
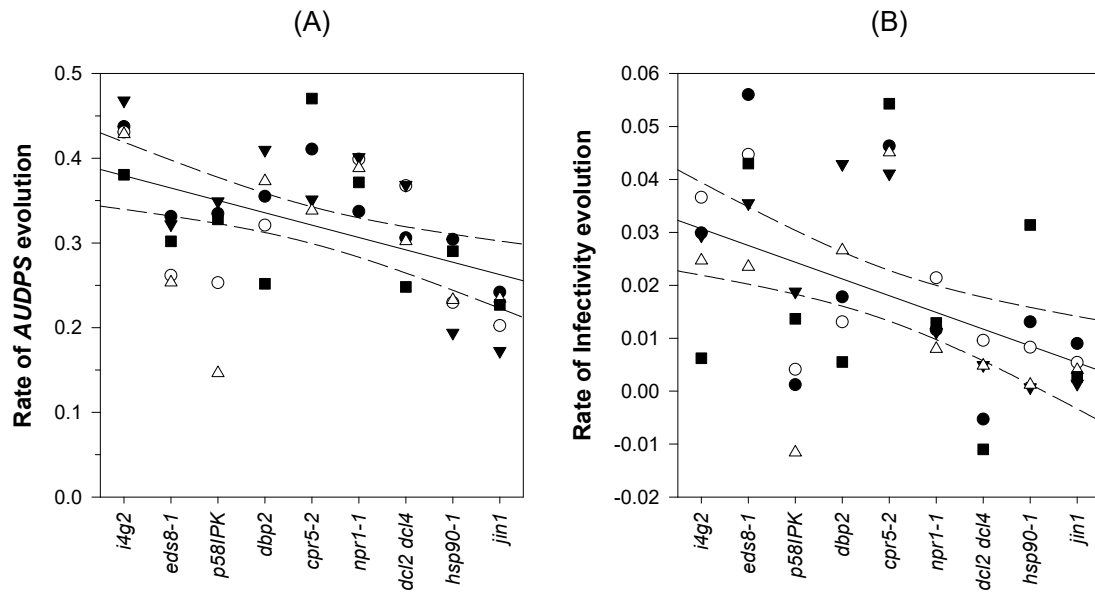
1147

Fig. 6. Estimated rates of evolution for *AUDPS* (A) and *I* (B) obtained from the fitting of an ARIMA(1,0,0) model (Eq. 2) to the data shown in Fig. 2 and Fig. 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, Arabidopsis genotypes are ranked from the less to the more permissive to TuMV infection according to Fig. 1.
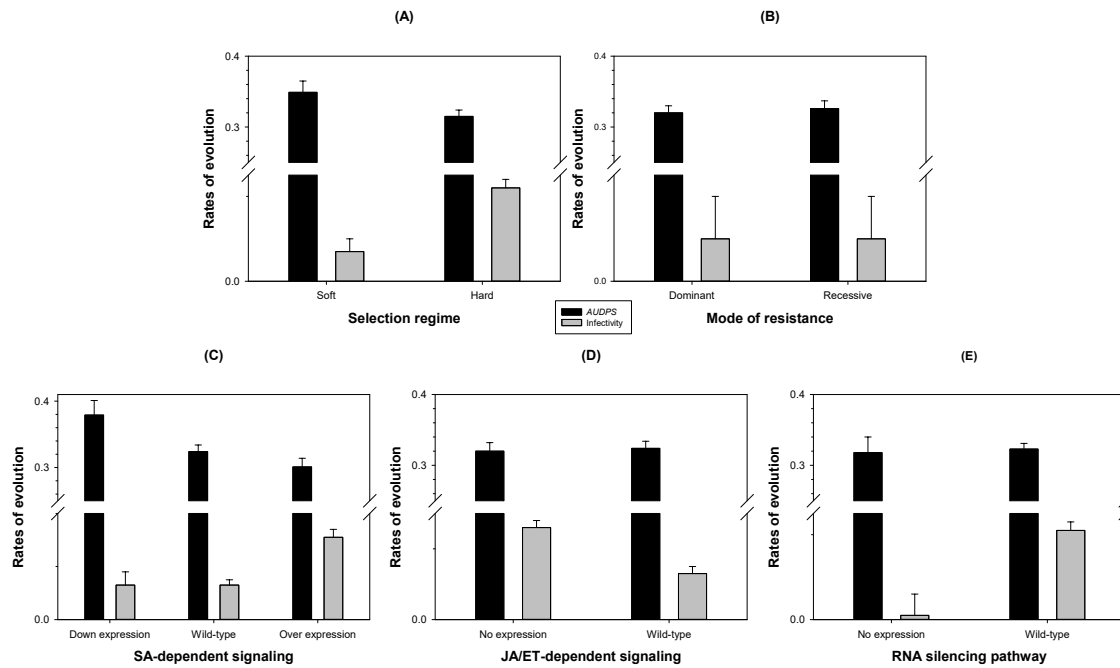
1148

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 Eq. 3. Error bars represent ±1 SEM.
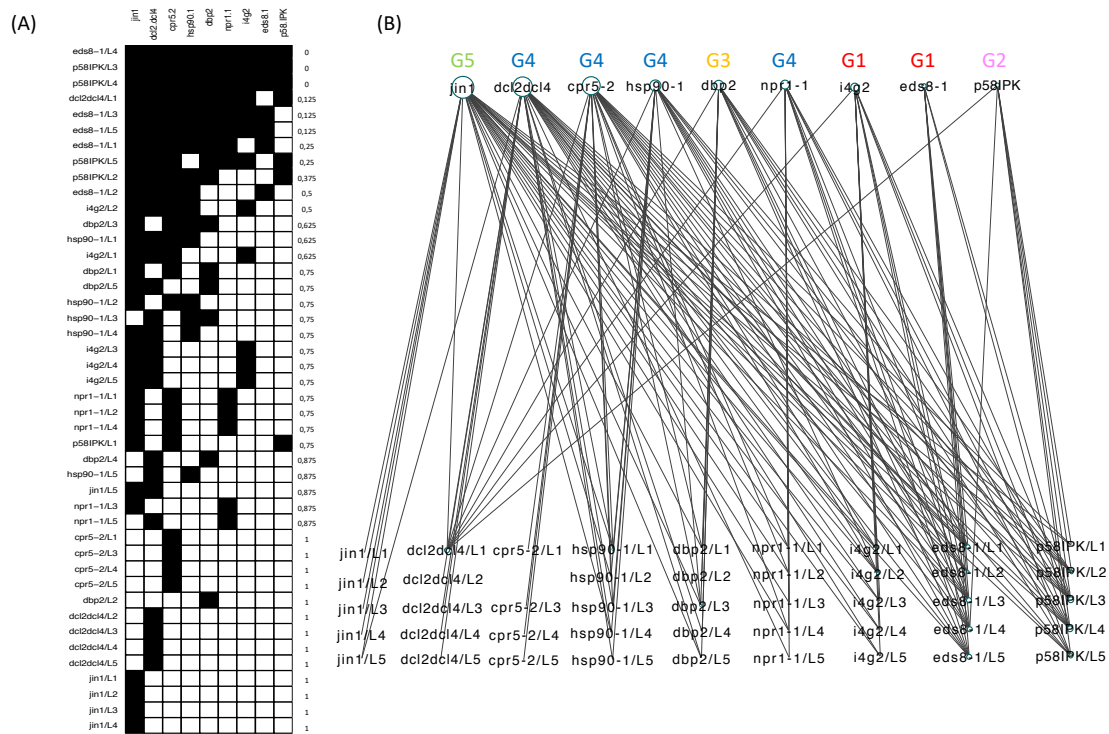
1149

Fig. 8. Analysis of the full cross-infection matrix. (A) Packed matrix that highlights the 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 Fig. 1 are indicated above the host genotype.
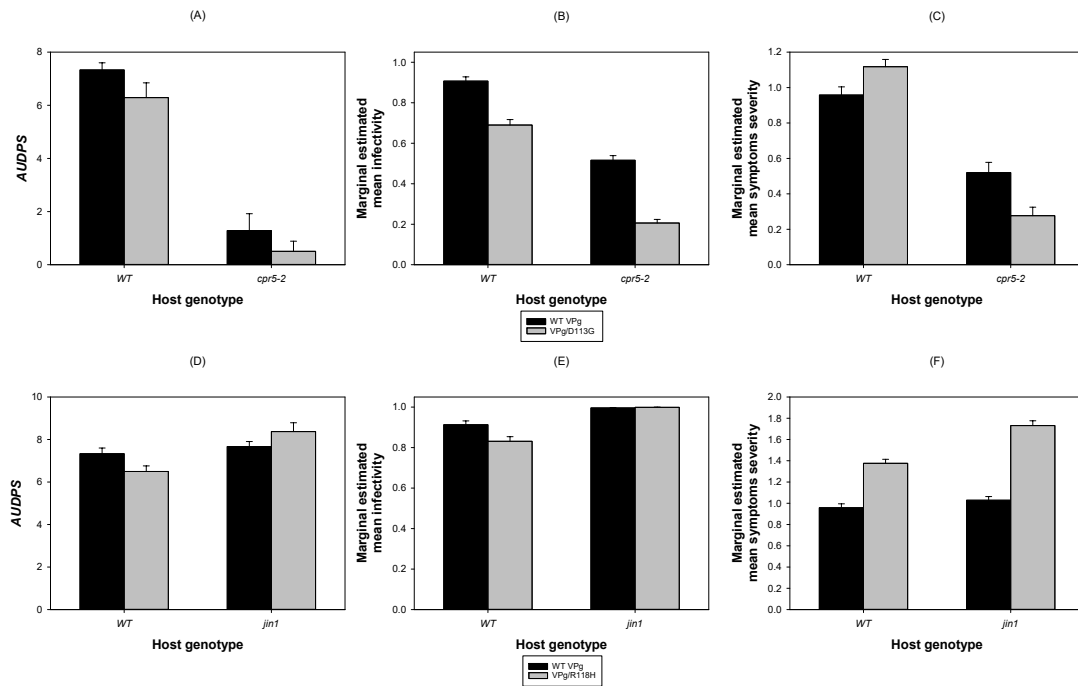
1150

54

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

1151

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

Genotype	Gene name	Affected pathway	Expected phenotype relative to wild-type plants	Reference
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

1152

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 Eq. 1.

Source of variation	Wilk's Λ	F	df_1, df_2	P	η_P^2	$1 - \beta$
μ	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.

1153

1154

Table 3. Results of the probit regression testing for the specificity of adaptation. The different model factors are defined in Eq. 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.
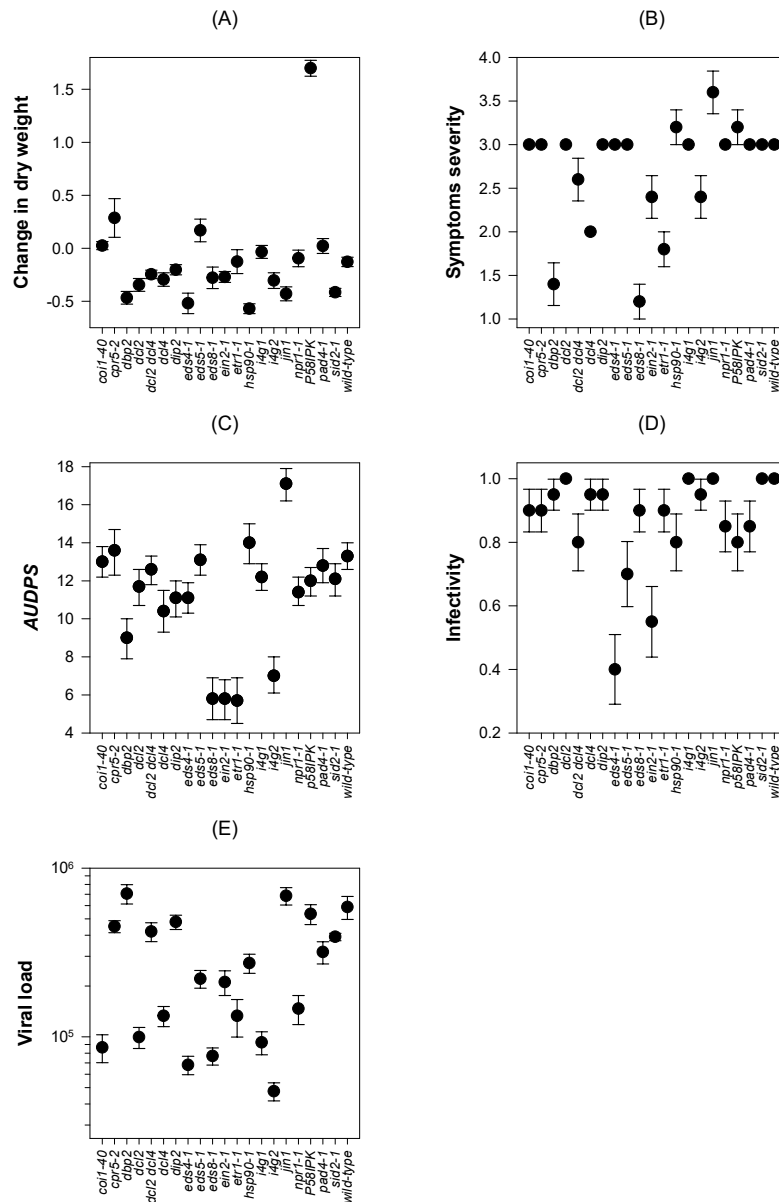
1155

Additional Information

Fig. S1. Disease-related traits evaluated for the 21 Arabidopsis genotypes listed in Table 1. (A) Change in dry weight (ΔDW; genetic component of variance $\sigma_G^2 = 93.09\%$, one-way ANOVA $F_{20,84} = 71.820$, $P < 0.001$). (B) Severity of symptoms (SS; $\sigma_G^2 = 75.88\%$, $F_{20,84} = 17.643$, $P < 0.001$). (C) Area under the disease progress stairs ($AUDPS$; $\sigma_G^2 = 99.96\%$, $F_{20,399} = 2626.754$, $P < 0.001$). (D) Infectivity (I; $\sigma_G^2 = 78.91\%$, $F_{20,399} = 4.563$, $P < 0.001$). (E) Viral load (VL; $\sigma_G^2 = 78.99\%$, $F_{20,84} = 22.113$, $P < 0.001$). Error bars represent ±1 SEM, except for the $AUDPS$, which are 95% confidence intervals based in 1000 bootstrap pseudo-replicates.
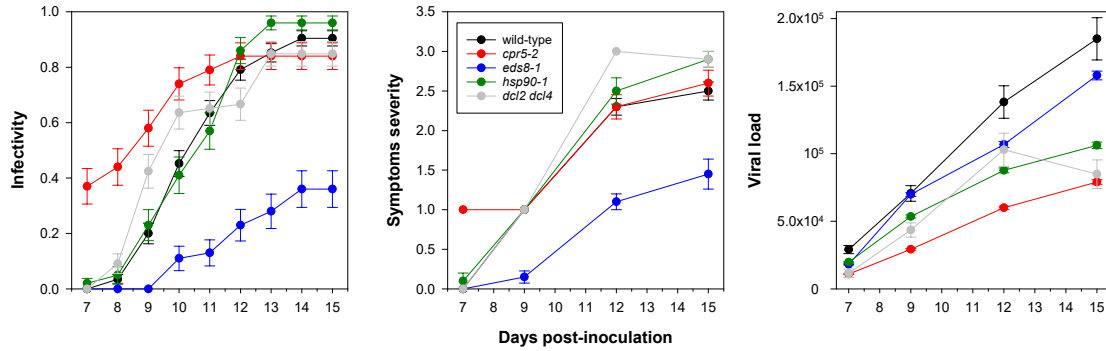
1159

Fig. S2. Daily variation of (A) infectivity (*I*), (B) symptoms severity (*SS*) and (C) viral load (*VL*) for genotypes representative of groups G1 (*eds8-1* more resistant than wild-type to TuMV infection) and G4 (*cpr5-2*, *dcl2 dcl4*, *hsp90-1* similar to wild-type in response to infection) in Fig. 1. Error bars represent ±1 SEM.
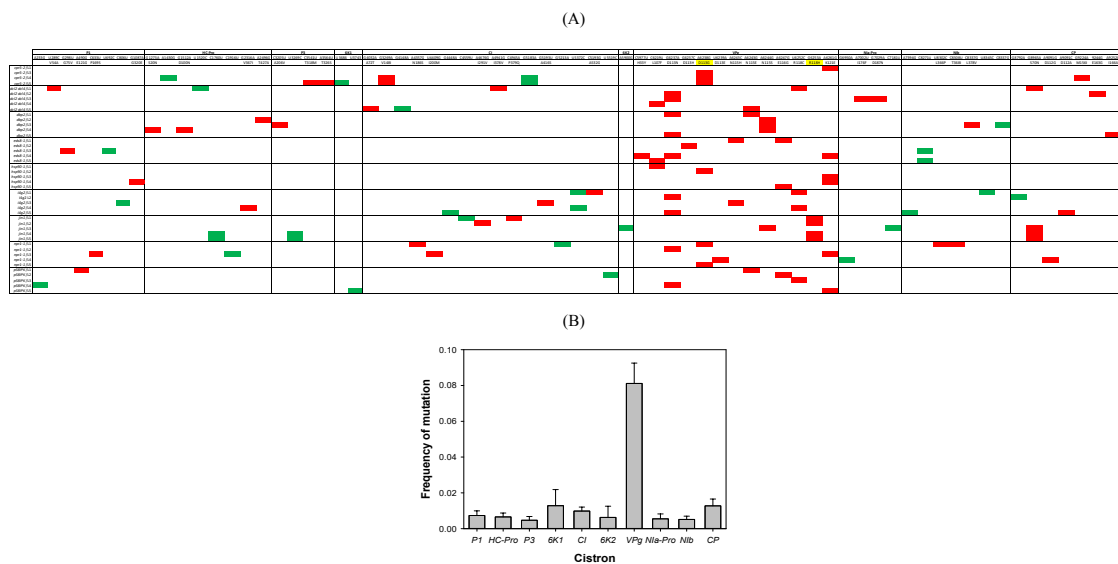
1160

Fig. S3. Distribution of observed mutations per TuMV cistron. (A) Schematic representation of the distribution of synonymous (green boxes) and nonsynonymous (red boxes) distribution along TuMV genome. (B) Frequency of mutations observed on each cistron. Error bars represent ±1 SEM.

1161

Table S1. Nucleotide and amino acid changes found in the consensus sequences of the TuMV evolved viruses. SNP positions referred to the complete genome. Amino acid changes positions are referred to the mature peptide.

ORF	Lineages	Mutation	Amino acid change	Amino acid group change	Syn:Nosyn	Ts:Tv
P1	p58^{IPK}/L4	A233G	synonymous			
	dcl2 dcl4/L1	U289C	V45A	conservative nonpolar by nonpolar		
	eds8-1/L3	G298U	G75V	conservative nonpolar by nonpolar		
	p58^{IPK}/L1	A490G	E121G	negatively charged by nonpolar		
	npr1-1/L3	C633U	P169S	nonpolar by nonpolar, strong structural effect	3:5	7:1
	eds8-1/L3	U692C	synonymous			
	i4g2/L3	C806U	synonymous			
	hsp90-1/L4	G1087A	G320E	nonpolar by negatively charged		
HC-Pro	dbp2/L4	G1273A	S20N	small polar by larger polar		
	cpr5-2/L4	A1430G	synonymous			
	dbp2/L4	G1512A	D100N	negatively charged by polar		
	dcl2 dcl4/L1	U1520C	synonymous		5:4	9:0
	jin1/L4 jin1/L5	C1760U	synonymous			
	npr1-1/L3	C1916U	synonymous			

	Gene/Line	Mutation	AA change	Description		
	i4g2/L4	G2316A	V367I	conservative small nonpolar by longer nonpolar		
	dbp2/L2	A2493G	T326S	small polar by nonpolar		
P3	*dbp2*/L3	C3205U	A206V	conservative nonpolar by nonpolar		
	jin1/L4 *jin1*/L5	U3269C	synonymous		2:3	3:1
	cpr5-2/L5	C3541U	T318M	small polar by long nonpolar		
	cpr5-2/L5	A3564U	T326S	conservative polar by polar		
6K1	cpr5-2/L5	U3686C	synonymous		2:0	2:0
	p58^{IPK}/L5	U3743C	synonymous			
	dcl2 dcl4/L5	G4032A	A72T	small nonpolar by polar		
	dcl2 dcl4/L5	G4148A	synonymous			
	cpr5-2/L4 *cpr5-2*/L5	G3269A	V148I	conservative small nonpolar by larger nonpolar		
CI	*npr1-1*/L1	A4357G	N186S	positively charged by small polar	6:13	15:4
	npr1-1/L3	U4409G	I200M	conservative nonpolar by nonpolar		
	i4g2/L5	G4448A	synonymous			
	jin1/L1	C4559U	synonymous			
	jin1/L2	A4674G	I291V	conservative nonpolar by nonpolar		
	dcl2 dcl4/L1	A4941G	I378V	conservative nonpolar by nonpolar		

	jin1/L1	C4945A	P379Q	nonpolar by polar, strong structural effect		
	cpr5-2/L4	G5183A	synonymous			
	cpr5-2/L5					
	i4g2/L3	G5193U	A416S	nonpolar by polar		
	npr1-1/L1	G5213A	synonymous			
	i4g2/L1	U5372C	synonymous			
	i4g2/L4					
	i4g2/L1	C5193G	A532G	conservative nonpolar by nonpolar		
	p58^{IPK}/L2	U5519C	synonymous			
6K2	*jin1*/L3	A5900G	synonymous		1:0	1:0
	eds8-1/L4	C5977U	H33Y	positively charged by polar		
	eds8-1/L5					
	hsp90-1/L1	C6219U	L107F	conservative nonpolar by nonpolar		
VPg	*dcl2 dcl4*/L4				0:47	45:2
	dbp2/L1					
	dbp2/L5	G6237A	D113N	negatively charged by polar		
	i4g2/L2					
	i4g2/L5					

dcl2 dcl4/L2			
dcl2 dcl4/L3			
eds8-1/L4			
p58IPK/L4			
npr1-1/L2			
eds8-1/L2	G6237C	D113H	negatively charged by positively charge
cpr5-2/L3			
cpr5-2/L4			
cpr5-2/L5	A6238G	D113G	negatively charged by nonpolar
npr1-1/L1			
npr1-1/L5			
hsp90-1/L2			
npr1-1/L4	U6239A	D113E	conservative negatively charged by negatively charged
i4g2/L3	A6243C	N115H	polar by positively charged
eds8-1/L3			
dbp2/L1			
dcl2 dcl4/L5	A6243G	N115E	polar by negatively charged
p58IPK/L1			
dbp2/L2	A6244G	N115S	conservative polar by polar
dbp2/L3			

dbp2/L4			
jin1/L3			
eds8-1/L1			
hsp90-1/L5	A6247G	E116G	negatively charged by small nonpolar
p58^{IPK}/L2			
i4g2/L1			
i4g2/L4	U6252C	R118C	positively charged by nonpolar
dcl2 dcl4/L1			
p58^{IPK}/L3			
jin1/L1			
jin1/L2			
jin1/L2	G6253A	R118H	conservative positively charged by positively charged
jin1/L4			
jin1/L5			
hsp90-1/L3			
hsp90-1/L4			
cpr5-2/L1	A6261G	K121E	positively charged by negatively charged
npr1-1/L3			
eds8-1/L4			
p58^{IPK}/L5			

NIa-Pro	*npr1-1*/L4	G6950A	synonymous			
	dcl2 dcl4/L3	A7002U	I176F	nonpolar by nonpolar		
	dcl2 dcl4/L3	G7029A	D187N	negatively charged by polar	2:2	3:1
	jin1/L3	C7181U	synonymous			
NIb	*i4g2*/L5	A7394G	synonymous			
	eds8-1/L3 *eds8-1*/L5	C8271U	synonymous			
	npr1-1/L5	U8302C	L366P	nonpolar by nonpolar, strong structural effect		
	npr1-1/L5	C8308U*	T368I	polar by nonpolar	5:3	7:1
	dbp2/L3	C8337G	L378V	conservative nonpolar by nonpolar		
	i4g2/L1	U8345C	synonymous			
	dbp2/L3	G8360A	synonymous			
CP	*hsp90-1*/L2	G8792A	synonymous			
	jin1/L3 *jin1*/L4 *jin1*/L5 *dcl2 dcl4*/L1	G8965A	S70N	conservative polar by polar	1:10	1:10
	npr1-1/L4	A9091G	D112G	negatively charged by small nonpolar		
	i4g2/L5	A9091C	D112A	negatively charged by small nonpolar		

cpr5-2/L4	G9224A	M156I	conservative nonpolar by nonpolar	1162
dcl2 dcl4/L2	A9244G	E163G		
dbp2/L5	A9252G	I166V	negatively charged by small nonpolar	
hsp90-1/L1	C9541U	A262V	conservative nonpolar by nonpolar	

[*]Polymorphic positions; Syn:Nosyn: number of synonymous *vs* nonsynonymous mutations; Ts:Tv: number of transitions *vs* transversions. The two highlighted nonsynonymous mutations D113G and R118H in VPg have been used in further experiments.