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

Rebeca Navarro¹, Silvia Ambrós¹, Fernando Martínez¹, Beilei Wu^{1,§}, José L. Carrasco¹, and Santiago F. Elena^{1,2,*}

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

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

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

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

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

Abstract

It is assumed that host genetic variability for susceptibility to infection will necessarily condition the evolution of viruses, either by driving them to diversification into strains that track the different host defense alleles (*e.g.*, antigenic diversity), or by canalization to infect only the most susceptible genotypes. Associated to these processes, virulence may or may not increase. To tackle these questions, we performed evolution experiments with turnip mosaic virus (TuMV) in *Arabidopsis thaliana* genotypes that differ in mutations in genes involved in resistance pathways and in genes whose products are essential for potyviruses infection. Plant genotypes classified into five groups according to their degree of resistance and intensity of symptoms. We found that evolution proceeded faster in the most resistant hosts than in the most permissive ones, as expected for adaptation to a harsh environment. The multifunctional viral protein VPg turned out to be the target of selection in most host genotypes. When all evolved TuMV lineages were tested for fitness in all plant genotypes used in the experiments, we found that the infection matrix was significantly nested, suggesting the evolution of generalist viruses selected by the most restrictive mutant genotypes. At the other side, a modular pattern, driven by convergent evolution of lineages evolved in the same host genotype, was also observed.

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 intricate signaling mechanisms (Zhou & Zhang 2020) that limit the spread of the virus, resulting in resistance. A variety of factors contribute to plant

resistance to viral infections. Broadly speaking, these factors can be classified into *basal* if they are pre-existing and limit within-cell propagation and cell-to-cell spread, and *inducible* if they are only activated upon infection and inhibit systemic virus replication and movement. Basal mechanisms include susceptibility (*S*) genes that involve alleles of cellular proteins that do not interact properly with viral factors, specially translation initiation factors, required by the virus for successful exploitation of the cell's protein synthesis machinery, heat shock proteins, that assist them in the formation of multiple multiprotein complexes, or DNA binding phosphatases (Truniger & Aranda 2009; Carr *et al.* 2010; Mäkinen 2019). In contrast, inducible mechanisms include most of resistance (*R*) genes whose expression result in a broad-scale change in plant physiology via diverse signal transduction pathways, particularly those regulated by the hormones salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) (Soosaar *et al.* 2005; Carr *et al.* 2010). These changes include local cell apoptosis (*e.g.*, hypersensitive responses - HR; Loebenstein 2009), the upregulation of nonspecific responses against many different types of pathogens throughout the entire plant (systemic acquired resistance – SAR- and induced systemic resistance – ISR) (Kachroo *et al.* 2006; Carr *et al.* 2010), and the activation of the RNA-silencing-based resistance, which seems to play a role both in basal and inducible mechanisms (Voinnet 2001; Carr *et al.* 2010). Early host responses following virus detection include changes in ion fluxes (mainly Ca⁺⁺), activation of signaling pathways, major alterations of transcriptomic profiles, generation of reactive oxygen species (ROS) and production of nitric oxide (NO) (Soosaar *et al.* 2005). These immediate changes are followed by HR and the recruitment of SA and JA/ET signaling pathways. The SA-mediated defense signaling pathway results in SAR, while the JA/ET-mediated defense signaling pathway results in ISR, the latter being specifically involved in symbiotic interactions between plants and beneficial microbes. Indeed, it appears that

ISR is not effective against infection by most viruses (Ton *et al.* 2002; Loebenstein 2009; Pieterse *et al.* 2009). Interestingly, both SAR and ISR pathways converge into two master regulators, the *ENHANCED DISEASE SUSCEPTIBILITY 1* (*EDS1*) and the *PHYTOALEXIN DEFICIENT 4* (*PAD4*) genes; *EDS1* and *PAD4* repress ISR and promote SAR. Although the SA and JA/ET pathways have been classically viewed as mutually antagonistic, several studies have revealed positive and negative crosstalk between them (van Wees *et al.* 2000; Pieterse *et al.* 2012) as well as between them and the RNA-silencing pathway (Soosaar *et al.* 2005; Carr *et al.* 2010; Yang *et al.* 2020). This crosstalk, which has the network topological structure of an incoherent feed-forward loop, creates robustness and tunability in the plant immune network (Mine *et al.* 2017).

A particularly well study example of *S* genes in the context of genetic variability for resistance and virus adaptation in *Arabidopsis thaliana* (L.) Heynh are the *RESTRICTION TO TEV MOVEMENT* (*RTM*) genes, composed by at least five different loci (Sofer *et al.* 2017) that encode for a variety of functions needed by tobacco etch virus (TEV; genus *Potyvirus*, family *Potyviridae*), and other potyviruses, to induce a systemic infection of the plant. The mechanism of resistance in this case is well understood: dominant alleles in all five loci result in resistance, whereas homozygous deletions in at least one of the loci result in increased susceptibility to infection and different degrees of systemic infection. By means of experimental evolution, it has been shown that TEV can readily adapt to different plant ecotypes. In particular, serial passages in susceptible plant genotypes *rtm1/rtm1* result in a viral isolate (TEV-*At17*) capable of infecting otherwise fully resistant genotypes (Agudelo-Romero *et al.* 2008; Lalić *et al.* 2010). Additional passages of TEV-*At17* in plants with different *RTM* configurations resulted in increased local adaptation and, more interestingly, in the evolution of a GFG-like interaction mechanism: the most susceptible plant genotype selected for less pathogenic and highly

specialized viruses while the most resistant plant genotype selected for more virulent and generalist TEV isolates (Hillung *et al.* 2015). Later on, these results were confirmed with turnip mosaic virus (TuMV; genus *Potyvirus*, family *Potyviridae*) in the context of fully monomorphic or maximally polymorphic experimental *A. thaliana* populations (González *et al.* 2019).

However, these previous studies have barely paid any attention to the intricacies of the underlying defense regulatory pathways of *A. thaliana*, the existence of multiple defense responses (*e.g.*, SAR, ISR, RNA-silencing, and the diversity of *S* genes) and the potential crosstalk among them. Therefore, an open question is how this intricacy determines the evolutionary fate of viruses. For instance, would two independent defense pathways select for specialist viruses adapted to counteract each one thus conforming a fitness tradeoff? Would a tradeoff in the investment of plant resources between pathways (*e.g.*, SAR vs ISR) result in a diversification of virus adaptive strategies? Would viruses target crosstalk points across pathways thus evolving as generalists? Here we will use the TuMV-*A. thaliana* experimental pathosystem to explore this sort of questions. After exploring the variability in phenotypic responses to TuMV infection of a collection of *A. thaliana* mutants in different *S* and *R* genes, we choose nine mutants which represent the entire diversity in phenotypic responses, including both basal and inducible resistance mechanisms. Then, we performed evolution experiments on each plant mutant genotype and track the evolution of several disease-related traits. At the end of the evolution experiment, we evaluated the effect of host genotypes in the rates of virus evolution and the contribution of contingency, selection and stochasticity into the outcome of evolution. Next, we explored whether the different TuMV lineages evolved as specialist or generalists depending on the mutated defense mechanism of their local hosts. Finally,

we sought to identify the molecular changes experienced in the genome of the different viral lineages and explored the possible adaptive value of a few convergent mutations.

Results

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

The 21 *A. thaliana* mutant genotypes used in the study are shown in Table 1, including information about the affected signaling pathways or cellular processes as well as the expected phenotype of infection relative to wild-type (WT) plants based on the description of these mutants. Genotypes were classified according to their phenotypic response to TuMV infection based in five different disease-related traits measured 18 days post-inoculation (dpi). Fig. 1 shows the neighbor-joining clustering of genotypes according to their multi-trait phenotypic response to TuMV infection (data in Fig. S1); we found five significant groups (hereafter named as G1 to G5). For all four members of G1 the response was highly homogenous and consistent with an enhanced resistance response to TuMV infection: no significant change in ΔDW whilst a significant and consistent reduction in *SS*, *AUDPS*, *I*, and *VL* relative to infected WT plants. Three members of G1 had enhanced SAR response (Table 1) and the fourth one (*i4g2*) is a well-known *S* gene involved in plant resistance to potyvirus (Nicaise *et al.* 2007; Gallois *et al.* 2010). The only member of G2 was the strong apoptosis-inducer mutant *p58^{IPK}*, which shows no significant changes in *SS*, *AUDPS*, *I*, and *VL* relative to infected WT plants but a significant reduction in ΔDW , indeed, infected plants were heavier than the noninfected controls, thus being this a case of increased tolerance to infection. The only member of G3 was mutant *dbp2*, also described as specifically increasing resistance against another

potyvirus (Castelló *et al.* 2011), which shows a very interesting response to TuMV infection: ΔDW was significantly increased while *SS*, *AUDPS* and *I* were significantly reduced; no effect in *VL* was observed. Thus, *dbp2* represents an intermediate response to TuMV infection between the highly resistant G1 and those genotypes in G4. G4 represents a sort of hotchpotch formed by genotypes that at most differ from WT plants in two traits, but without a clear pattern. Remarkably, though, with the exception of *cpr5-2*, *dcl2 dcl4*, and *dip2*, all show significant reductions in *VL*. Therefore, members from G4 would be considered similar to WT in their phenotypic response to TuMV infection. Finally, *jin1* was the only member of G5, which shows a significant increase in all measured traits except *VL*, consistent with and enhanced susceptibility to TuMV infection. This is a surprising finding since *jin1* has been described as inducing a constitutive expression of the SA-dependent defenses that made plants extremely resistant to the infection of biotrophic bacterial pathogens (Laurie-Berry *et al.* 2006). To further characterize the progression of infection, we selected a subset of representatives from groups G1 (*eds8-1*) and G4 (*cpr5-2*, *dcl2 dcl4* and *hsp90-1*) and evaluated daily *I*, *SS* and *VL* (Fig. S2). The *cpr5-2* plants started showing symptoms faster, though only 80% of plants showed symptoms after 15 dpi. Only 20% of plants of genotype *eds8-1* showed symptoms of infection, which were delayed on time. Infected *hsp90-1* and *dcl2 dcl4* plants were undistinguishable from the WT infected ones. Regarding *VL*, all genotypes accumulated less virus than WT plants, with *cpr5-2* being the one that did so the less (Fig. S2)

Representatives from the five groups were selected for the subsequent evolution experiment, covering the entire spectrum of responses to TuMV infection and mechanisms of resistance/susceptibility (shadowed rows in Table 1): *eds8-1* and *i4g2* from G1 representing the hardest selection regimes (more resistant genotypes); *p58^{IPK}*

from G2; *dbp2* from G3; *cpr5-2*, *dcl2 dcl4*, *hsp90-1*, and *npr1-1* from G4; and *jin1* from G5 representing the softest selective regime (less resistant genotype).

Experimental evolution in the representative *A. thaliana* genotypes

First, we sought to evaluate the dynamics of evolution for *AUDPS* (Fig. 2), *I* (Fig. 3) and *VL* (Fig. 4). Data in these figures were fitted all together to the multivariate analysis of covariance (MANCOVA) model described in Eq. 1 of the Methods section. Notice that lineage *cpr5-2/L2* showed a quite different evolution pattern from the other four lineages evolved in *cpr5-2*. It never increased in either of the three traits and was lost after passage five (Fig. 2, Fig. 3). Therefore, we ended up with 44 evolved lineages.

Table 2 shows the results of this MANCOVA. Despite a considerable amount of noise in the time series, we found some significant results that can be summarized as follows. First, a net significant and large ($\eta_p^2 = 0.735$) effect associated to the evolutionary passages has been observed, indicating that the three phenotypic traits related to TuMV infection evolved during the experiment. Second, plant genotypes had a highly significant and large ($\eta_p^2 = 0.547$) effect on the phenotypic traits, suggesting that TuMV evolutionary dynamics were strongly influenced by the local host genotype in which it was being passaged. More interestingly, this effect associated to host plant genotypes was strongly dependent on the passage number ($\eta_p^2 = 0.463$), thus suggesting that the slope of the relationship between phenotypic traits and time was influenced by the local host genotype. This effect will be further evaluated in the next section. Third, independent lineages show a strong degree of evolutionary parallelism, as indicated by nonsignificant differences among viral lineages evolved in the same local host genotype (factors $L(G)$ and $t \times L(G)$). However, this last conclusion should be taken with certain

level of caution since the power of the two tests is relatively low ($1 - \beta < 0.800$) and hence we might be wrongly failing to reject the null hypothesis (type II error).

A relevant question in evolutionary biology is the extent in which ancestral differences determine the fate of evolution. In other words, the relative contribution of adaptation, chance and historical contingency to the evolution or organismal fitness. Following the logic exposed by Travisano *et al.* (1995), we can imagine two situations. First, ancestral statistical differences among phenotypes could be preserved along evolution despite a net increase in the mean trait values (due to selection) and differences among replicated lineages (due to chance; *i.e.*, mutation and drift). In this situation, we should expect a nonzero slope in a regression of the evolved phenotypic values against the ancestral ones. The closer the slope to a value of one, the more importance of ancestral differences. Second, if initial trait variation among ancestral genotypes was eliminated from the evolved populations because the combined effect of adaptation and chance, we should expect a regression slope smaller than one. The less effect of ancestral differences, the flatter the slope, being zero in the extreme case of ancestral differences being completely erased. To evaluate the role of the ancestral differences in *AUDPS*, *I* and *VL* of TuMV across the nine different plant genotypes, we evaluated the magnitude of their change at the end of the evolution experiment. Fig. 5 shows the plots of evolved vs ancestral values. For illustrative purposes, the diagonals represent the null hypothesis of absolute preservation of ancestral differences. In the case of *AUDPS* (Fig. 5A), a significant regression exists between evolved and ancestral values ($R = 0.510$, $F_{1,42} = 14.732$, $P < 0.001$), with a slope of 0.296 ± 0.077 (± 1 SD). Since the slope is still significantly different from zero, yet clearly flatter than the diagonal ($t_{42} = 9.112$, $P < 0.001$), we conclude that ancestral differences have been mostly removed by the combined action of selection and chance yet not completely erased. In the case of *I* (Fig.

5B) the regression was not significant ($R = 0.087$, $F_{1,43} = 0.031$, $P = 0.568$), thus suggesting that for this trait ancestral differences among *A. thaliana* genotypes have been fully erased by adaptation and chance. Finally, in the case of *VL* (Fig. 5C), the regression of evolved values on ancestral ones was significant ($R = 0.417$, $F_{1,43} = 8.822$, $P = 0.005$), suggesting again that ancestral phenotypic differences have been largely removed by selection and chance yet some still persist (regression slope -0.142 ± 0.048 significantly different from one: $t_{42} = 23.903$, $P < 0.001$).

Rates of phenotypic evolution for disease-related traits

As mentioned in the previous section, we have observed a significant effect of *A. thaliana* genotypes in the temporal evolution of the three phenotypic traits. To further explore this effect, we estimated the rates of phenotypic evolution of *AUDPS* and *I* using the ARIMA(1,0,0) model described in the Methods section (Eq. 2). *VL* was not included in this analysis since the number of data points (three) in the time series was the same than the number of parameters to be estimated, thus likely incurring in overfitting problems.

Fig. 6 shows the estimated rates of evolution, β , for host genotypes ordered from the most resistant (G1) to the most permissive (G5) ones. Very interestingly, a significant negative regression coefficient exists for both traits [*AUDPS*: $-1.460 \cdot 10^{-2} \pm 0.412 \cdot 10^{-2}$ ($R = 0.480$, $F_{1,42} = 12.561$, $P = 0.001$) and *I*: $-3.161 \cdot 10^{-3} \pm 0.909 \cdot 10^{-3}$ ($R = 0.473$, $F_{1,42} = 12.105$, $P = 0.001$)], suggesting that evolution always proceeded faster in the most restrictive hosts and slower in the most permissive ones.

Dissecting the role of specific defense mechanisms in TuMV evolution

The rates of TuMV phenotypic evolution estimated in the previous section were further analyzed in the context of the mutated defense signaling pathway encoded by *R* genes or

recessive resistances encoded by *S* genes. Rates of evolution β_{AUDPS} and β_I were fitted to the MANOVA model shown in Eq. 3 of the Methods section. Fig. 7 graphically summarizes the results of these analyses. First, we sought for differences in rates of evolution between permissive (soft selection) and restrictive (hard selection) hosts (Fig. 7A). The multivariate analysis shows that lineages evolved in the more restrictive hosts evolved significantly faster than those evolved in the more permissive hosts ($\Lambda = 0.327$, $F_{2,34} = 34.923$, $P < 0.001$), with the effect being mostly driven by differences in the rate of infectivity evolution β_I , as shown by the corresponding univariate analyses, being on average 3.14 times faster in the most restrictive hosts. This observation confirms the results shown in the previous section.

Next, we evaluated whether differences exist between the rates of phenotypic evolution for viruses evolving in host genotypes carrying mutations affecting resistance (*R*; *i.e.*, the SAR, ISR and RNA-silencing pathways) or susceptibility (*S*; *dbp2*, *dip2*, *hsp90-1*, *i4g2*, and *p58^{IPK}*) genes (Fig. 7B). In this case, the multivariate analysis found a small yet significant effect ($\Lambda = 0.790$, $F_{2,34} = 4.507$, $P = 0.018$), though none of the two univariate analyses had enough power to detect it (indeed, changes went in opposite directions: 1.02-fold larger for β_{AUDPS} in the *S* plants but 1.31 times larger for β_I in the *R* plants).

Fig. 7C shows the comparison of rates of TuMV phenotypic evolution for lineages evolving in plant genotypes mutant for the SA-dependent signaling defense pathways. In this case we have three categories: mutations inducing a down-expression of the pathway, mutations with no effect on this pathway and mutations resulting in overexpression. Again, the MANOVA found highly significant differences among the three categories ($\Lambda = 0.213$, $F_{4,68} = 19.861$, $P < 0.001$). The univariate analyses show that most of the differences were due to a faster evolution of *I*, being 2.38 times faster in the host

genotypes over-expressing the SAR defenses than in those with a WT or even down-expression of the SA signaling.

SA- and JA/ET-dependent pathways are considered as antagonistic defenses, although more recent evidences point towards an active crosstalk between both (van Wees *et al.* 2000; Pieterse *et al.* 2012; Mine *et al.* 2017). Under the antagonistic hypothesis, we should expect the opposite trend than observed for the lineages evolved in plants over-expressing SAR defenses (Fig. 7C). Fig. 7D shows the evolution rates data classified according to whether host genotypes had mutations affecting the JA/ET defense (ISR) signaling pathway. The MANOVA found a significant difference between the two categories ($\Lambda = 0.433$, $F_{2,34} = 22.258$, $P < 0.001$), entirely driven by differences in β_I , as confirmed by the univariate analyses. Indeed, as expected, the rates of evolution for infectivity were 2-fold faster for TuMV lineages evolved in the ISR deficient hosts, in agreement with the SAR/ISR antagonistic hypothesis.

Finally, we tested the effect of knocking down the host's RNA-silencing pathway in the rates of TuMV evolution (Fig. 7E). In this case, the multivariate analysis also found highly significant differences between rates of viral evolution in plants with a fully functional RNA-silencing pathway and *dcl2 dcl4* plants ($\Lambda = 0.490$, $F_{2,34} = 17.729$, $P < 0.001$), with average β_I values being 21-fold larger in plants expressing a fully functional RNA-silencing pathway than in *dcl2 dcl4* plants unable of producing siRNAs.

In conclusion from this section, we found that regardless the defense signaling pathway, viruses evolving in permissive plants always evolve slower than their counterparts evolving in more restrictive plant genotypes.

Permissiveness of *A. thaliana* genotypes to infection drive the evolution of specialist or generalist TuMV lineages

To analyze the specificity of adaptation of each evolved TuMV lineage, we performed a full cross-infection experiment in which all the 44 evolved lineages were inoculated into ten plants of the nine *A. thaliana* genotypes used in the evolution experiments. The presence of symptoms on each plant was evaluated daily for up to 12 dpi. These data were analyzed using two different approaches, first an ANOVA-like that evaluates the effect of host genotypes, viral lineages and their interactions (Schmid-Hempel, 2011); and second, the inference and characterization of nestedness, modularity and specialization of an infection network (Weitz *et al.* 2013). For the first approach we fitted the presence/absence of symptoms data to the logistic regression model shown in Eq. 4 of the Methods section. The results are shown in Table 3. Despite its apparent complexity, interpretation is straightforward. A *TG* factor describes the effect of differences among host genotypes, irrespective of the infecting viral lineage; it thus characterizes whether some host genotypes are more susceptible than other to infection. As shown in Table 3, a *TG* has a highly significant effect, both by itself and in the interaction with *t*. The magnitude of the net effect was very large ($\eta_P^2 = 0.569$), while its interaction with *t* was rather small in magnitude ($\eta_P^2 = 0.010$). On average, the most susceptible host genotypes were *jin1*, *dcl2 dcl4* and *cpr5-2* ($\bar{f} = 1.000$), whereas the most resistant host genotypes were *p58^{IPK}* ($\bar{f} = 0.459 \pm 0.013$) and *eds8-1* ($\bar{f} = 0.233 \pm 0.013$). Similarly, a highly significant *LG* effect exists by itself (of large magnitude $\eta_P^2 = 0.288$) as well as in combination with *t* (though the magnitude of the interaction was rather small in magnitude $\eta_P^2 = 0.005$), which means that the observation depends on the viral lineage independently of the host genotype. On average, the lineages evolved in *eds8-1* ($\bar{f} = 0.997 \pm 23.710$) and *dbp2* ($\bar{f} = 0.994 \pm 26.108$) were those inducing symptoms faster; by contrast, lineages evolved in *i4g2* ($\bar{f} = 0.788 \pm 279.090$) showed symptoms much slower. Finally, and most interesting, a highly significant interaction term *TG*×*LG* (as well as in

combination with t) indicates that the outcome depends on the particular combinations of host genotypes and TuMV lineages. Indeed, the magnitude of the interaction effect was large ($\eta_P^2 = 0.314$), though its dependence with dpi was not so relevant ($\eta_P^2 = 0.048$). For example, in the extreme of the most susceptible genotypes, the case of *cpr5-2* perfectly illustrates this case: lineages evolved in *cpr5-2*, *dbp2*, *dcl2 dcl4*, *eds8-1*, and *p58^{IPK}* all show $\bar{f} = 1.000$, whereas lineages evolved in *i4g2* and *jin1* show much lower $\bar{f}$ values (0.674 ± 0.028 and 0.849 ± 0.032 , respectively). In the opposite extreme, e.g. in *eds8-1* lineages evolved in *eds8-1* show $\bar{f} = 0.575 \pm 0.050$ while lineages evolved in *jin1* show $\bar{f} = 0.124 \pm 0.030$; all other lineages showing intermediate values.

For the second analytical approach, we used the above infection dataset to estimate the corresponding *AUDPS* values, thus creating an infection matrix of 44 rows [$L(LG)$] by nine columns (TG). This matrix was transformed into a new binary matrix in which a value of one meant that the *AUDPS* for a given TuMV lineage in a particular TG was greater or equal than observed for its LG and zero otherwise. Fig. 8A shows the packed infection matrix with rows and columns organized to better highlight its nestedness and modularity (Weitz *et al.* 2013). The matrix could be used to build a binary infection network (Fig. 8B) that provides the same information in a more graphical way. The infection network was significantly nested ($T = 14.706$, $P < 0.001$), with the most permissive genotypes *jin1* and *dcl2 dcl4* being susceptible to most viral lineages (universally susceptible hosts), while the most resistant genotype *eds8-1* and *p58^{IPK}* were only successfully infected by some lineages that were precisely evolved in these two genotypes. Likewise, the most generalist viral lineages were *eds8-1/L4*, *p58^{IPK}/L3* and *p58^{IPK}/L4* that infected with equal efficiency all nine host genotypes (universally virulent viruses), while the most specialized viral lineages, able only to successfully infect their local hosts, were all surviving lineages evolved in *cpr5-2*, four lineages evolved in *dcl2*

dcl4 and four lineages evolved in *jnl1* (Fig. 8). Therefore, we conclude that local host's resistance to infection positively correlates with the host range of the evolved lineages: more resistant host genotypes selected for very generalist viruses while less resistant host genotypes selected for very specialized viruses, in agreement with a gene-for-gene infection model (Weitz *et al.* 2013).

Finally, we computed Blüthgen *et al.* (2006) specialization indexes for the packed matrix shown in Fig. 8. Firstly, we evaluated the standardized species-level measure of partner diversity d' . d' ranged from zero for the most generalist TuMV lineages (*eds8-1/L4*, *p58^{IPK}/L3* and *p58^{IPK}/L4*) to one for the most specialist ones (the four surviving *cpr5-2*-evolved lineages, *dbp2/L2*, *dcl2 dcl4/L2 - L5* and *jnl1/L1 - L4*) (last column in Fig. 8A). Secondly, we computed the network-level specialization index H'_2 , obtaining a value of $H'_2 = 0$ (uncorrected $H_2 = 5.024$, $H'_{min} = 3.534$, $H'_{max} = 5.024$; $P < 0.001$), which means that, overall, the binary infection network shows a degree of specialization, with most lineages being able of infecting one or few host genotype, a result mainly driven by those TuMV lineages evolved in the less resistant host genotypes.

Finally, we sought to explore if the infection network in Fig. 8B also shows evidences of modularity. Indeed, from an ecological standpoint, a significant network-level degree of specialization means that some TuMV lineages do not interact well with some of the *A. thaliana* genotypes, thus creating the possibility of modularity. In this context, a module will refer to an aggregated set of viral lineages characterized by more interactions within the module than between modules (Newman 2006; Dormann & Strauss 2014). To this end, we computed the Q modularity index, which ranges from zero, when the community has no more links within modules than expected by chance, to a maximum value of one. We found a small yet significant modularity ($Q = 0.234$, $P = 0.047$) in the infection matrix (Fig. 8A). What are the implications of modularity in

TuMV pathogenesis? Modularity is expected for a MA infection model (Weitz *et al.* 2013). More similar host genotypes, *e.g.*, those with mutations affecting the same signaling pathway, may be selecting for viral lineages with similar properties, thus being able of infecting with equal efficiency the subset of plant genotypes. A less appealing alternative hypothesis to explain the observed modularity may be parallel evolution of lineages evolved into the same host genotype; the most representative cases of this possibility being the lineages evolved in *cpr5-2*, *dcl2 dcl4* and *jin1*. To distinguish between these two hypotheses, we computed a reduced matrix averaging the observed *AUDPS* among TuMV lineages evolved in the same host genotype and transforming this nine by nine infection matrix into a packed binary matrix as described above. If the reduced matrix still shows significant modularity, the first hypothesis will hold, while if modularity disappears by averaging lineages evolved in the same host genotype, then the alternative hypothesis will be more parsimonious. The reduced matrix is still significantly nested ($T = 13.064$, $P < 0.001$), but modularity was not significant anymore ($Q = 0.286$, $P = 0.137$), thus supporting the alternative hypothesis that the observed modularity was driven by convergent evolution of lineages evolved into the same host genotype rather than by the overlap of selective pressures by genotypes carrying mutations in the same defense mechanism.

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

Finally, we have explored the molecular changes experienced by the TuMV evolved lineages. Full genomic consensus sequences were determined for all 44 evolved lineages. A total of 114 mutational events have been observed, affecting 71 nucleotide position (Supplementary Fig. S3A and Table S1). According to the type of nucleotide substitution

involved, 93 were transitions and 21 transversions; regarding their effect on the protein sequence, 30 were synonymous and 84 nonsynonymous. Interestingly, some mutations, including 13 nonsynonymous and four synonymous, have been observed multiple times in independent lineages. As discussed in the Methods section, treating each lineage as an observation and each host genotype as a subpopulation, the average nucleotide diversity within host genotypes, referred only to the 71 polymorphic sites, is $\pi_S = 0.067 \pm 0.007$. On the other hand, the nucleotide diversity for the entire sample is $\pi_T = 0.074 \pm 0.007$. Consequently, the estimate of interhost genotypes nucleotide diversity is $\delta_{ST} = 0.007 \pm 0.003$, and the estimated coefficient of nucleotide differentiation (Nei 1982), $N_{ST} = 0.090 \pm 0.033$, a value significantly greater than zero ($z = 2.727$, $P = 0.003$). Thus, we conclude that minor yet significant genetic differentiation has been generated among viral lineages replicating in different host genotypes. To assess whether selection played a role in genetic differentiation among *A. thaliana* genotypes, we performed a Tajima's *D* test (Tajima 1989) and found that it was significantly negative ($D = -2.487$, $P = 0.006$).

Next, we sought to characterize the distribution of mutational events along the nine non-overlapping cistrons (Fig. S3B). The frequency of mutations per cistron, relative to the length of the corresponding cistron, was fitted to the logistic regression model shown in Eq. 5 of the Methods. Highly significant differences exist ($\chi^2 = 127.545$, 9 d.f., $P < 0.001$), yet entirely due to the ~16-fold larger mutation frequency observed in the *VPg* cistron relative to the rest of the genome (Fig. S3B). Notice that all mutations observed in *VPg* are nonsynonymous and that all lineages except *eds8-1/L3* carry at least one mutation in this cistron (Fig. S3A and Table S1).

Convergent nonsynonymous mutations are, *a priori*, good candidates for adaptive mutations. Mutation CP/V148I appears in two lineages evolved in *cpr5-2* (Fig. S3A and Table S1), mutation CP/S70N appear in lineage *dcl2 dcl4/L1* and three lineages evolved

in *jin1* (Fig. S3A and Table S1), and two different mutations affecting CP amino acid 112, CP/D112G (lineage *npr1-1/L4*) and CP/112A (lineage *i4g2/L5*), that result in a similar replacement of side chains. In addition, seven nonsynonymous mutations in VPg are shared by several lineages (Fig. S3A and Table S1). Out of these seven cases, three seem particularly promising candidates. They all are located in a narrow region of VPg (residues 113, 115 and 118) (Fig. S3A and Table S1). Firstly, mutations G6237A, G6237C, A6238G, and U6239A all affect the same codon, resulting in amino acid replacements D113N, D113H, D113G, and D113E, respectively. The expected functional effect of these mutations was evaluated using the algorithms implemented in SNAP2 (Hecht *et al.* 2015). D113E is predicted to be functionally neutral (score = -76, accuracy = 87%), while D113G (32, 66%), D113H (43, 71%) and D113N (30, 66%) are predicted to have a functional effect. Among these three amino acid replacements, D113G, which shows an intermediate value of the functional effect score, is particularly interesting, involving a very strong change in the side radical from a long negatively charged one to a small nonpolar one. This replacement has been observed in lineages *cpr5-2/L3*, *cpr5-2/L4*, *cpr5-2/L5*, *npr1-1/L1*, *npr1-1/L5*, and *hsp90-1/L2*. All three hosts for these lineages belong to the phenotypic group G4 in Fig. 1. Secondly, mutations A6243C, A6243G and A6244G also affect the same codon, resulting in amino acid replacements N115H (-42, 72%), N115E (-66, 82%) and N115S (-30, 61%). All three are predicted to be functionally neutral. Thirdly, mutations U6252 and G6253A affect the same codon and result in amino acid replacements R118C (-10, 53%) and R118H (56, 78%), respectively. The most frequent among these two replacements is R118H, which represents a conservative change among large positively charged side radicals, yet has a strong expected functional effect. R118H has been observed in four out of five TuMV lineages evolved in the *A. thaliana* mutant *jin1* (phenotypic group G5 in Fig. 1).

D113G thus represents a case of convergent evolution non-specific of the local host genotype (*i.e.*, a candidate for a generalist mutation) while R118H represents a mutation highly specific of the local host genotype (*i.e.*, a specialist mutation).

To further characterize the possible adaptive value of mutations VPg/D113G and VPg/R118H, we created by site-directed mutagenesis the two mutated versions of *VPg* and cloned them into the ancestral p35STunos infectious clone (Chen *et al.* 2003). Viruses were recovered from these clones and three disease-related phenotypic traits (*e.g.*, *AUDPS*, *I* and *SS*) evaluated both in the WT plants as well as in their corresponding local hosts (*i.e.*, *cpr5-2* for mutant VPg/D113G and *jin1* for mutant VPg/R118H). Fig. 9 summarizes the results of these experiments. Regarding VPg/D113G (Fig. 9A - C), the results strongly depend on the combination of phenotypic trait and plant genotype: it shows a significant negative effect in both *AUDPS* (−14.2%; $z = 1.687$, $P = 0.046$) and *I* (−23.9%; $z = 6.344$, $P < 0.001$) but a positive one in *SS* (16.7%; $z = 2.597$, $P = 0.005$) when evaluated in WT plants. However, it shows a largely negative yet not significant effect in *cpr5-2* plants for *AUDPS* (−60.8%; $z = 1.055$, $P = 0.146$) and significant negative effects both for *I* (−60.1%; $z = 10.614$, $P < 0.001$) and *SS* (−46.7%; $z = 3.228$, $P = 0.001$). These negative effects on the local host genotype *cpr5-2* do not support a possible beneficial effect of this mutation by itself. Regarding the possible adaptive effect of the amino acid replacement VPg/R118H (Fig. 9D - F), the results are also dependent on the combination of phenotypic trait and host plant genotype. When tested in WT plants, the mutation has significant negative effects both for *AUDPS* (−11.4%; $z = 2.29$, $P = 0.013$) and *I* (−9.0%; $z = 2.749$, $P = 0.003$) but a significantly positive effect for *SS* (43.6%; $z = 7.881$, $P < 0.001$). By contrast, in the local host *jin1* the effect was positive in all three traits; significant for *AUDPS* (9.2%; $z = 1.683$, $P = 0.046$) and *SS* (68.2%; $z = 12.447$, $P < 0.001$) but not for *I* (0.3%; $z = 1.342$, $P = 0.090$). These results support the conclusion

that the amino acid replacement VPg/R118H has an overall beneficial effect on the local host genotype *jin1* in which it was selected.

Discussion

Interaction between TuMV and the different defense pathways

In this study we have explored the effect that mutations in host's different disease signaling pathways (*R* genes) or in recessive resistance genes (*S*) have in the outcome of virus evolution. Our experimental pathosystem consisted in two well studied components: the model plant *A. thaliana* and one virus with high prevalence in natural populations on this host (Pagán *et al.* 2010), TuMV. Among the *R* genes, we have included in our study genes involved in the SA, JA/ET and RNA-silencing pathways, among the *S* genes, we have included heat shock proteins, transcription factors and components of the translation machinery. In a preliminary set of experiments, we established the similarities and differences in phenotypic responses of different plant mutant genotypes to TuMV infection. Our results were, in some cases, at odds with the expected phenotypes for *R* genes, being the results for *jin1* the most unexpected ones. *JASMONATE INSENSITIVE 1 (JIN1)* is a negative regulator of SA-mediated defense responses, henceforth the mutant *jin1* has a constitutive expression of SAR (Laurie-Berry *et al.* 2006). Surprisingly, it turns out to be the most sensitive genotype to TuMV infection, showing enhanced symptoms of disease (phenotypic group G5). This may reflect the fact that the function of *R* genes, especially those involved in SAR and ISR, have been mostly defined in terms of plant interactions with biotrophic and necrotrophic bacteria and fungi but rarely in response to virus infections. Mutant genotypes only affecting components of the ISR pathway (*e.g.*, *coil-4*) behaved as wild-type plants in

response to TuMV infection, thus confirming previous reports that ISR was inefficient against viral infections (Ton et al. 2002; Loebenstein 2009; Pieterse *et al.* 2009). However, mutant *eds8-1*, which avoids ISR but enhances SAR (Love *et al.* 2007) turned out to be among the most resistant genotypes to infection (phenotypic group G1). In agreement with our findings of a variable response to SA-signaling, Singh *et al.* (2004) discussed examples of viruses in which SA-dependent responses to viral infection were strongly dependent on the virus used in the experiments: while viruses such as alfalfa mosaic virus, potato virus X or turnip vein clearing virus replication was inhibited by treating plants with SA, cucumber mosaic virus was apparently unaffected by the treatment and accumulated to normal levels.

The RNA-silencing pathway is considered as the main plant defense against viral pathogens (Voinnet 2001), with *DICER LIKE 2* (*DCL2*) and *DICER LIKE 4* (*DCL4*) encoding for the two dicer enzymes responsible for generating the 22- and 21-nucleotides long antiviral siRNAs. Therefore, the double mutant *dcl2 dcl4* was expected to show a highly sensitive response to TuMV infection. However, infected plants were, overall, hardly distinguishable from the WT plants in their response to infection (Fig. S2 and Fig. S3). A possible explanation would be the strong suppressor activity of HC-Pro in the WT plants that effectively counteracts the defense mechanism (Kaschau & Carrington 2001). Indeed, it has been shown that the outcome of the interplay between plant RNA-silencing response and potyvirus is mostly driven by the efficiency of the viral HC-Pro suppression activity (Li *et al.* 2007; Torres-Barceló *et al.* 2008).

The response to infection of genotypes carrying mutations in *S* genes was, in general, consistent with the *a priori* expectation. Mutant *i4g2*, that is defective for the eIF4(iso)E factor, shows a strong resistance to TuMV infection (Nicaise *et al.* 2007; Charron *et al.* 2008), belonging to phenotypic group G1. Likewise, in agreement with

previous descriptions, mutant *dbp2* shows enhanced resistance to potyvirus infection (Castelló *et al.* 2011), belonging to phenotypic group G3. Finally, mutant *hsp90-1* turns out to show a response to TuMV infection equivalent to the WT plants, likely due to its functional redundancy with other heat-shock proteins present in the cell that may be coopted by the virus to assist in their expected functions (Verchot 2012).

An interesting finding was the effect on TuMV replication of knocking down the expression of gen *P58^{IPK}*. In mammals *P58^{IPK}*, a tetratricopeptide repeat (TPR)-containing protein, is recruited by viruses (*e.g.*, influenza A virus) to inhibit interferon activation and cell death mediated by the dsRNA-activated protein kinase (PKR), thus favoring viral spread (Goodman *et al.* 2011). In our experiments, *p58^{IPK}* mutant plants show enhanced tolerance to infection, showing weaker symptoms than WT plants despite accumulating viruses at the same level (Fig. S1). In sharp contrast, Bilgin *et al.* (2003) found that infection of *p58^{IPK}*-silenced *Nicotiana benthamiana* and *A. thaliana* plants infected with tobacco mosaic virus lead to death. Given the limited information available from other plant-virus pathosystems, additional experiments would be necessary to shed light into the potential antiviral role in plants of this gene.

One side, yet important, conclusion of our study is that the generally accepted model of plant defense signaling pathways, mainly based in experiments done with bacteria and fungi, may not well describe the interaction between plants and viruses. By contrast, recessive resistances based on *S* genes seem to better explain differences in susceptibility to infection and viral accumulation, thus being more promising targets for future development of resistant plants.

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

The GFG and MA models of host-virus interaction modes represent the two ends of a continuum of possible outcomes (Agrawal & Lively 2002). The most relevant difference between both models regards the expected genetic heterogeneity in both host and virus populations. With a pure GFG interaction the susceptible host types are expected to disappear and the resistant types will dominate the population. *Vice versa*, the most virulent virus allele would drive to fixation at the cost of losing mild alleles. However, constitutive activation of defenses is known to be costly for *A. thaliana* [e.g., SA-related defense responses pay a fitness cost in absence of pathogens (Traw *et al.* 2007)] and high virulence usually comes with a cost in terms of pathogen's transmission (reviewed in Acevedo *et al.* 2019). Hence, a pure GFG strategy seems unlikely to be achieved. By contrast, with a pure MA interaction, negative frequency-dependent selection emerges, such that rare *A. thaliana* resistance alleles have advantage and, as a result, a genetic polymorphism shall be maintained (Schmid-Hempel 2011). Natural populations of *A. thaliana* contain considerable amount of genetic variability for tolerance (Pagán *et al.* 2008) and for immunity-related genes (Todesco *et al.* 2010, van de Weyer *et al.* 2019), thus suggesting that an arms race between pathogens and plants should be going on yet it is unclear whether it may result in pure GFG or MA interactions or lie somewhere in between. Evolution experiments with such different pathosystems as TuMV/*A. thaliana* (González *et al.* 2019), *Octospora bayeri*/*Daphnia magna* (Altermatt & Ebert 2008) and *Serratia marcescens*/*Caenorhabditis elegans* (Gibson *et al.* 2020; White *et al.* 2020) have produced congruent results in one aspect: parasites exposed to heterogeneous host populations evolved significantly lower virulence than parasites exposed to homogeneous host populations. However, a significant difference exists among pathosystems: while viruses exposed to genetically heterogeneous host populations evolved as no-cost

generalists, evolution of generalism in more complex parasites was constrained by a fitness tradeoff, as expected for the jack-of-all trades hypothesis (Bedhomme *et al.* 2015).

Our results, as well as those by Hillung *et al.* (2014) and González *et al.* (2019) have shown the evolution of significantly nested binary infection networks, a finding compatible with the existence of a combination of specialist and generalist viruses and of more permissive and resistant host genotypes. Indeed, these studies have also shown that more permissive hosts selected for more specialized viruses while more resistant hosts selected for more generalist viruses, here matching the predictions of the GFG model. Our observation of small yet significant modularity in the infection network was easily explained by convergent evolution of TuMV lineages evolved in the same host genotype. However, it has been recently shown in a long-term survey of the prevalence of different plant viruses in different hosts and habitats that nestedness and modularity in host-pathogen infection networks is possible due to the spatially patched distribution of habitats and temporal successions of plant species (Valverde *et al.* 2020): small spatial scales create modularity that coexist with global nestedness. This pattern may change spatially and temporally but remains stable over long evolutionary timescales.

Role of natural selection

We found evidences of significant genetic differentiation among TuMV lineages evolved in different plant genotypes. To test whether these differences were driven by selection we performed a Tajima's *D* test (Tajima 1989). The resulting *D* value was significative and negative, which is compatible with the action of purifying selection, the presence of slightly deleterious mutations segregating in the populations or fast population expansions (Yang 2006). How to distinguish between these explanations? In independent fast expanding populations, many new mutations may be generated and

rising in frequency in each population, thus being observed as singletons, mutations present in only one of the many coexisting genomes in each evolving TuMV lineage. Singletons inflate the number of segregating sites and thus cause $D < 0$. Indeed, this is the case here: 54 out of the 71 observed variable sites are singletons, thus the observed pattern of molecular diversity among lineages evolved in the same host genotype and in different host genotypes is likely to be due to the fast expansions of viral populations.

However, we have found additional evidence supporting the action of positive selection: the existence of a number of convergent nonsynonymous mutations arising in independent lineages, some evolved in the same host genotype but some others rising in different host genotypes (Table S1 and Fig. S3) yet without a clear association with the particular signaling pathway or *S* gene being mutated (Table 1) nor the phenotypic cluster they belong to (Fig. 1). Interestingly, most of these convergent mutations happened in the *VPg* cistron, which turns out to be also the most variable one. *VPg* plays many essential roles in genome transcription (it is linked to the 5'-end of the viral genome and provides the hydroxyl group that primes the synthesis of the complementary strands by the viral RdRp), translation (directly interacts with the eukaryotic initiation factors *eIF(iso)4E* and *eIF(iso)4G*) and interacts with all other viral proteins (Bosque *et al.* 2014) and some other cell proteins (Martínez *et al.* 2016). Indeed, in previous evolution experiments with potyviruses, *VPg* has also been shown to be an important target of selection. For example, Agudelo-Romero *et al.* (2008) found that a single amino acid replacement in *VPg* was enough to largely increase TEV *I*, *VL* and *SS* in *A. thaliana*. Similarly, Gallois *et al.* (2010) found that *A. thaliana* plants with knock-out mutations in the *eIF(iso)4E*, *eIF(iso)4G1* and *eIF(iso)4G2* genes were resistant to TuMV infection. Two mutations in the *VPg* (E116Q and N163Y) were enough to overcome this resistance and return to the original infection phenotype, though yeast-two hybrid assays showed

that none of these mutations affected the binding of VPg with eIF(iso)4E. As a final example, one of the most extensively used resistance genes in pepper commercial cultivars against potato virus Y (PVY; genus *Potyvirus*, family *Potyviridae*) is the *pvr2*, which has many different alleles (Nicaise *et al.* 2007; Charron *et al.* 2008). The *pvr2* locus encodes for the eIF4E factor which, as mentioned above, physically interacts with VPg. Interestingly, all the resistance-breaking viral isolates found so far contain mutations in the *VPg* cistron (Duprat *et al.* 2002; Moury *et al.* 2004; Ayme *et al.* 2006)

Interestingly, one of the two mutations identified by Gallois *et al.* (2010), VPg/E116Q, affects the same protein domain that the two mutations we have identified and characterized (D113G and R118H). The mechanism why these two mutations may confer a selective advantage to TuMV lineages cannot be inferred from our studies. We found VPg/D113G in several lineages evolved in different host genotypes, while VPg/R118H was only found in the *jin1* lineages. Our assays failed to find a beneficial effect of VPg/D113G but confirmed the host-specific beneficial effect of VPg/R118H in *jin1* plants.

Unfortunately, due to economical constraints, we have not been able to exhaustively testing all observed mutations and, hence, several others may still be potential candidates for adaptive mutations. More specifically, we have not tested possible epistatic interactions among mutations fixed in the same genome. Three lineages carry more than one nonsynonymous mutation in *VPg* (*dbp2*/L1 D113N and N115E, *eds8-1*/L1 N115H and E116G, and *eds8-1*/L4 H33Y, D113N and K121E) (Table S1). In addition, 25 lineages carry additional nonsynonymous mutations in at least one other cistron besides *VPg* (Table S1). Two lineages have up to three additional nonsynonymous mutations: *cpr5-2*/L5 that has P3/T318M-T326S CI/V148I and VPg/D113G and *dcl2 dcl4*/L1 that has P1/V54A CI/I378V VPg/R118C and CP/S70N. Therefore, plenty of opportunities

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

Methods

Plants, virus and growth conditions

A collection of 21 different *A. thaliana* mutants of the Col-0 accession were used for this study (Table 1). In all experiments described below, plants were all maintained in a BSL-2 climatic chamber under a photoperiod of 8 h light (LED tubes at PAR 90 - 100 $\mu\text{mol}/\text{m}^2/\text{s}$) at 25 °C and 16 h dark at 20 °C.

Prior to the inoculation experiments, we created a large stock of TuMV infectious saps. Saps were obtained from TuMV-infected *N. benthamiana* Domin plants inoculated with the infectious plasmid p35STunos that contains a cDNA of TuMV genome (GeneBank accession AF530055.2) under the control of the cauliflower mosaic virus 35S promoter and the *nos* terminator (Chen *et al.* 2003) as described elsewhere (González *et al.* 2019; Corrêa *et al.* 2020). This TuMV sequence variant corresponds to the YC5 isolate from calla lily (*Zantedeschia* sp) (Chen *et al.* 2003). After plants showed symptoms of infection, they were pooled frozen with liquid N₂ and homogenized into a fine powder using a Mixer Mill MM400 (Retsch GmbH, Haan, Germany). For inoculations, the necessary amount of powder was diluted in inoculation buffer (50 mM phosphate buffer pH 7.0, 3% PEG6000, 10% Carborundum) and 5 μL of the inoculum was gently rubbed into the plant leaves. Plants were all inoculated when reached growth stage 3.5 in the Boyes' scale (Boyes 2001). This synchronization ensures that they all were at the same phenological state when inoculated.

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

harvested and used to prepare infectious saps as described above. These saps were 1/10 diluted in inoculation buffer and used to inoculate the next batch of plants.

Total RNA extractions

Tissue from each pool of infected plants per lineage, ecotype and serial passage was collected, frozen with liquid N₂ and preserved at -80 °C until it was homogenized into fine powder using a Mixer Mill MM400. Next, an aliquot of approximately 100 mg of grounded tissue per sample for which viral load was quantified, was used for total RNA (RNA_{At}) extraction performed with the Agilent Plant RNA isolation Mini kit (Agilent Technologies, Santa Clara CA, USA). Aliquots of RNA_{At} per each sample were separated and their concentration adjusted at 50 ng/μL to estimate viral accumulation by RT-qPCR (see below).

Quantification of VL

VL of each plant sample per lineage, ecotype and selected passage was quantified by absolute real-time quantitative RT-PCR (RT-qPCR) using standard curves and the primers TuMV F117 forward (5'-CAATACGTGCGAGAGAAGCACAC-3') and F118 reverse (5'-TAACCCCTTAACGCCAAGTAAG-3') that amplify a 173 nucleotides fragment from the CP cistron of TuMV genome, as previously described (Corrêa *et al.* 2020). Briefly, standard curves were constructed using ten serial dilutions of the TuMV genome, that was synthesized by *in vitro* transcription as detailed previously (Cervera *et al.* 2018), in RNA_{At} extract obtained from healthy *A. thaliana* plants used as control in the experiments. Amplification reactions were run in a 20 μL volume using the GoTaq 1-Step RT-qPCR System (Promega, Madison WI, USA) and the recommended manufacturer's instructions as described (Cervera *et al.* 2018) in an ABI StepOne Plus

Real-time PCR System (Applied Biosystems, Foster City CA, USA). The cycling conditions consisted in: an RT phase of 5 min at 42 °C and 10 min at 95 °C followed by a PCR stage consisting in 40 cycles of 5 s at 95 °C and 34 s at 60 °C; and the final melt curve profile analysis that consisted in 15 s at 95 °C, 1 min at 60 °C and 15 s at 95 °C. Negative controls consisted in healthy RNA plant extract (non infected control) and water. Quantitative reactions were run as three technical replicates per sample and results were analysed using the StepOne software 2.2.2 (Applied Biosystems).

TuMV genome amplifications

Evolved viral genomes of the passage 12 from each lineage and ecotype were amplified by high-fidelity RT-PCR by using the AccuScript Hi-Fi (Agilent Technologies) reverse transcriptase and Phusion DNA polymerase (Thermo Scientific, Waltham MA, USA) following the manufacturer's instructions. Each complete TuMV genome was amplified into three overlapping amplicons of 3114 (5' fragment R1), 3697 (central region R2) and 3287 nucleotides (3' fragment R3) using three primer sets. For RT reactions an aliquot of the corresponding RNA (150-300 ng) was mixed with 0.25 µM of the 1R-P3 (5'-CGAGTAGTATCTTATAGCACAGCGCTCCGACC-3'), 2R-N1a (5'-TGTCTGGAATCGGTAGCAAATGTAGCTGAGTTGTG-3') or 3R-polyAR (5'-TTTTTTTTTTTTTTTTTTTTGTCCCTTGCATCATATCAAATG-3') primer to synthesize the R1, R2 or R3 cDNA fragment, respectively, that were denatured 5 min at 65 °C and cooled on ice. Then it was added a mix containing AccuScript Hi-Fi 1× Buffer, 1 mM of dNTPs, 8 mM of DTT, 4U of Ribolock RNase inhibitor (Thermo Scientific) and 0.5 µL of AccuScript Hi-Fi (Agilent Technologies) in a 10 µL volume. RT reactions consisted in a step of 90 min at 42 °C to synthesize the cDNA followed by an incubation of 5 min at 70 °C to inactivate the enzyme. PCR reactions were performed in a 50 µL

volume containing a mix of 1× Phusion Buffer, 0.4 μM of dNTPs, 0.2 μM of each primer, 0.5-1 μL of DMSO, 2U of Phusion DNA polymerase (Thermo Scientific) and 1 μL of the corresponding RT reaction. R1 fragment was amplified using the primer set 1F-5UTR (5'-GCAAACGCAGACCTTTCGAAGCACTCAAGC-3') and 1R-P3 and the following PCR conditions: an initial denaturation of 30 s at 98 °C, 3 cycles of 10 s at 98 °C, 20 s at 67 °C and 2 min at 72 °C, 3 cycles of 10 s at 98 °C, 20 s at 65 °C and 2 min at 72 °C, and 32 cycles of 10 s at 98 °C, 20 s at 63 °C and 2 min at 72 °C, followed by a final extension step of 5 min at 72 °C. Fragments R2 and R3 were amplified with primer set 2F-P3 (5'-TGGGAGCTTGCGGATGGTGGATACACAATTC-3') and 2R-NIa or 3F-NIa (5'-CTCGTTATATGGAGTCGGTTTCGGACCACTCATAT-3') and 3R-polyAR, respectively and a PCR with the same denaturation and extension steps than the fragment R1 but different amplification steps: an stage consisting in 15 cycles of 10 s at 98 °C, 20 s at 67 °C and 2 min at 72 °C followed by 23 cycles of 10 s at 98 °C, 20 s at 65 °C and 2 min at 72 °C for fragment R2, while 15 cycles of 10 s at 98 °C, 20 s at 67 °C and 2 min at 72 °C followed by 23 cycles of 10 s at 98 °C, 20 s at 65 °C and 2 min at 72 °C for fragment R3. PCR products were purified with the MSB Spin PCRapace Kit (Stratagene Molecular, Coronado CA, USA) and then Sanger-sequenced. Full-length consensus viral sequences were obtained assembling the sequences of the three amplified products by using the Genious R9.0.2 program.

Construction of TuMV *VPg* mutants

Mutants in the *VPg* cistron were synthesized by long inverse site-directed PCR mutagenesis using the QuickChange II XL Kit (Stratagene, San Diego CA, USA) following the manufacturer's instructions using p35STunos as template. Also, mutagenic primers were designed following the manufacturer's recommendations. Primer set cpr5-

2F (5'-GGAGGATGAGTTGGGTCCAAATGAAATACGTGT-3') and cpr5-2R (5'-ACACGTATTTCATTGGACCCAATCATCCTCC-3') was used to generate the A6238G (D113G) mutant while primer set jin1F (5'-GGATCCAAATGAAATACATGTGAATAAGACAATTC-3') and jin1R (5'-GAATTGTCTTATTCACATGTATTTCATTGGATCC-3') to obtain the G6253A (R118H) mutant. PCR protocol consisted in: a denaturation step of 2 min at 95 °C followed by 20 cycles of 20 s at 95 °C, 10 s at 60 °C and 7 min at 68 °C and a final extension of 5 min at 68 °C. After *DpnI* digestion and transformation of electrocompetent *Escherichia coli* DH5 α , the presence of the desired mutations in the infectious clone and the absence of undesired nucleotide changes was confirmed by sequencing.

Bioassays of the TuMV VPg mutants

The infectivity, viability and pathogenicity of both VPg TuMV mutants was confirmed by performing three independent bioassays. In each experiment, three batches of 24 *A. thaliana* plants (three-weeks old) of genotypes wild-type, *cpr5-2* and *jin1* were inoculated with the wild-type TuMV plasmid clone used as reference and with plasmids of each of the TuMV VPg mutants, while another two batches of 24 plants from *cpr5-2* or *jin1* genotypes were inoculated with the WT TuMV plasmid and the corresponding plasmid of the TuMV VPg mutant, the VPg/D113G mutant in the case of the *cpr5-2* ecotype and that of VPg/R118H for the *jin1* genotype. *A. thaliana* were inoculated by abrasion of three leaves applying equal amounts of the plasmid inoculum, a total of approximately 7 μ g per plant distributed in 3 μ L per leaf. Plasmids were previously purified by Midiprep using the NucleoBond Xtra Midi Kit (Macherey-Nagel, Düren, Germany) and resuspended in distilled water. Number of infected plants and symptom intensity were collected every day until 15 dpi.

Statistical analyses

All statistical analyses described hereafter were performed with SPSS version 26 software (IBM, Armonk, NY), unless otherwise indicated.

AUDPS (Fig. 2), *I* (Fig. 3) and *VL* (Fig. 4) data were fitted all together to a multivariate analysis of covariance (MANCOVA) model in which plant genotype (*G*) was the main factor, the independent evolution lineages (*L*) were nested within *G* and passage (*t*) was introduced in the model as a covariable. The full model equation thus reads:

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

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

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

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

where *Y_k* represents the variable being analyzed at passage *k*, ρ_1 measures the degree of self-similarity in the time-series data (correlation between values at passages *t* and *t* − 1),

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

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

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

Analysis of infection the network

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

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

where $f_{ijkl}(t)$ is the frequency of symptomatic plants t dpi in the test host genotype TG , for the viral lineage L evolved in the local host genotype LG . TG and LG were considered as orthogonal factors, L was nested within LG and dpi (t) was treated as a covariable. The

model includes all main factors, their corresponding nested and factorial interactions as well as their interactions with the covariable. ϕ represents the grand mean of the probit-transformed f values, and ε the error term assumed to be Binomial. The significance of the different factors was evaluated using likelihood-ratio tests (LRT).

For the second statistical approach, the binary infection matrix was analyzed using tools borrowed from the field of network biology to explore whether they show random associations between viral lineages and host genotypes, one-to-one associations, nestedness indicative of a GFG type of interaction, or modularity (Weitz *et al.* 2013). The statistical properties of the infection matrix were evaluated using the R package “bipartite” version 2.15 (Dormann *et al.* 2008) in R version 4.0.0 (R Core Team 2020) under RStudio 1.2.1335. Four different summary statistics were evaluated: T nestedness (Bascompte *et al.* 2003), Q modularity (Newman 2006) and the d' species-level (or Kullback-Leibler divergence) and H'_2 network-level (or two-dimensional normalized Shannon entropy) specialization indexes (Blüthgen *et al.* 2006). H'_2 ranges between zero and one for extreme generalists and specialists, respectively. Statistical significance of these statistics was evaluated using Bascompte *et al.* (2003) null model.

The statistics of molecular evolution

Treating each lineage as an observation and each host genotype as a subpopulation, we evaluated the average nucleotide diversity within host genotypes, π_S , the nucleotide diversity for the entire sample, π_T , the interhost genotypes nucleotide diversity, δ_{ST} , and the estimate of the proportion of interhost genotypes nucleotide diversity, known as coefficient of nucleotide differentiation (Nei 1982), $N_{ST} = \delta_{ST}/\pi_T$. Standard deviations of estimates were inferred from 1000 bootstrap samples. All these computations were done using MEGA X (Kumar *et al.* 2018) and the lowest-BIC nucleotide substitution model

Kimura 2-parameters (Kimura 1980). Tajima's D test of selection (Tajima 1989) and its statistical significance were evaluated using DnaSP6 (Rozas *et al.* 2017).

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

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

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

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

References

- Acevedo MA, Dilleuth FP, Flick AJ, Faldyn MJ, Elder BD. 2019. Virulence-driven trade-offs in disease transmission: a meta-analysis. *Evolution* **73**: 636-647.
- Agrawal A, Lively CM. 2002. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evol. Ecol. Res.* **4**: 79-90.
- Agudelo-Romero P, Carbonell P, Pérez-Amador MA, Elena SF. 2008. Virus adaptation by manipulation of host's gene expression. *PLoS ONE* **3**: e2397.
- Altermatt F, Ebert D. 2008. Genetic diversity of *Daphnia magna* populations enhances resistance to parasites. *Ecol. Lett.* **11**: 918-928.
- Altizer S, Kaitala V, Laakso J, Ruokolainen L. 2006. Seasonality and the dynamics of infectious diseases. *Ecol. Lett.* **9**: 467-484.

- Anttila J, 2015. Environmental variation generates environmental opportunist pathogen outbreaks. *PLoS ONE* **10**: e0145511.
- Ayme V, Souche S, Caranta C, Jacquemond M, Chadoeuf J, Palloix A, Moury B (2006) Different mutations in the genome-linked protein VPg of *Potato virus Y* confer virulence on the *pvr2³* resistance in pepper. *Mol. Plant Microbe Interact.* **19**: 557-563.
- Bascompte J, Jordano P, Melián CJ, Olesen JM. 2003. The nested assembly of plant-animal mutualistic networks. *Proc. Natl. Acad. Sci. USA* **100**: 9389-9387.
- Bedhomme S, Hillung J, Elena SF. 2015. Emerging viruses: why are not jacks of all trades? *Curr Opin Virol* **10**: 1-6.
- Bilgin DD, Liu Y, Schiff M, Dinesh-Kumar SP. 2003. *P58^{IPK}*, a plant ortholog of double-stranded RNA-dependent protein kinase PKR inhibitor, functions in viral pathogenesis. *Dev. Cell* **4**: 651-661.
- Blüthgen N, Menzel F, Blüthgen N. 2006. Measuring specialization in species interaction networks. *BMC Ecol.* **6**: 9.
- Bosque G, Folch-Fortuny A, Picó J, Ferrer A, Elena SF. 2014. Topology analysis and visualization of *Potyvirus* protein-protein interaction network. *BMC Syst. Biol.* **8**: 129.
- Bouché N, Laressergues D, Gascioli V, Vaucheret J. 2006. An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J.* **25**: 3347-3356.
- Boyes DC. 2001. Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *Plant Cell* **13**: 1499-1510.
- Brown JKM, Tellier A. 2011. Plant-parasite coevolution: bridging the gap between 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, Nicolai 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, Parinthewong 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. 2008. 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, Eppele 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, Montañez 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

1128 **Acknowledgments**

1129 We thank Francisca de la Iglesia and Paula Agudo for excellent technical assistance and
1130 the rest of the EvolSysVir lab members for fruitful discussions. This work was supported
1131 by grant PID2019-103998GB-I00 (Agencia Estatal de Investigación - FEDER) and
1132 PROMETEU2019/012 (Generalitat Valenciana) to S.F.E.

1133

1134 **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

1138 **Competing interests**

1139 The authors declare no competing interests.

1140

Figures and Tables

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.

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

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

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
<i>coil-4</i>	<i>CORONATINE INSENSITIVE 1</i> (AT2G39940)	Repression of JA-responsive genes	No ISR, no effect on virus	Thines <i>et al.</i> (2007)
<i>cpr5-2</i>	<i>CONSTITUTIVE EXPRESSOR OF PR GENES 5</i> (AT5G64930)	Membrane protein, negative regulator of pathogen-dependent SA signaling	More resistant, constitutive SAR	Love <i>et al.</i> (2007)
<i>dbp2</i>	<i>DNA-BINDING PROTEIN PHOSPHATASE 2</i>	Transcriptional regulation of gene expression in potyvirus-infected plants	More resistant to potyvirus infection	Castelló <i>et al.</i> (2011)
<i>dcl2</i>	<i>DICER-LIKE 2</i> (AT3G03300)	Partial loss of RNA-silencing	No effect, siRNAs produced by DCL4	Bouché <i>et al.</i> (2006)
<i>dcl4</i>	<i>DICER-LIKE 4</i> (AT5G20320)	Partial loss of RNA-silencing	No effect, siRNA produced by DCL2	Bouché <i>et al.</i> (2006)
<i>dcl2 dcl4</i>	Double mutant <i>dcl2 dcl4</i>	Complete loss of RNA-silencing	More susceptible, no siRNA production	Bouché <i>et al.</i> (2006)
<i>dip2</i>	<i>DBP-INTERACTING PROTEIN 2</i> (AT5G03210)	Transcriptional regulation of gene expression in potyvirus-infected plants	More susceptible	Castelló <i>et al.</i> (2011)
<i>eds4-1</i>	<i>ENHANCED DISEASE SUSCEPTIBILITY 4</i> (AT5G51200)	Loss of SA-dependent signaling	More susceptible, no SAR	Gupta <i>et al.</i> (2000)
<i>eds5-1</i>	<i>ENHANCED DISEASE SUSCEPTIBILITY 5</i> (AT4G39030)	Lipase-like protein, positive regulator of pathogen-dependent SA signaling	More susceptible, no SAR	Nawrath <i>et al.</i> (2002)
<i>eds8-1</i>	<i>ENHANCED DISEASE SUSCEPTIBILITY 8</i>	Reduced expression of plant defensin genes, reduced ISR	More resistant, enhanced SAR	Love <i>et al.</i> (2007)
<i>ein2-1</i>	<i>ETHYLENE INSENSITIVE 2</i> (AT5G03280)	MAPK, ET signaling intermediate, negative regulator SA-dependent signaling	More resistant, enhanced SAR	Love <i>et al.</i> (2007)
<i>etr1-1</i>	<i>ETHYLENE RESPONSE 1</i> (AT1G66340)	ET receptor, negative regulator SA-dependent signaling	More resistant, enhanced SAR	Love <i>et al.</i> (2007)

<i>hsp90-1</i>	<i>HEAT SHOCK PROTEIN 1</i> (AT5G52640)			Recessive <i>r</i> gene, required for membrane-bound replication complexes; protein folding	More resistant, missing component for viral replication	Verchot (2012)
<i>i4g1</i>	<i>EUKARYOTIC TRANSLATION INITIATION FACTOR (ISO) 4G 1</i> (AT3G60240)			Recessive <i>r</i> gene, initiation of viral RNA translation	More resistant, missing component for viral gene expression	Nicaise <i>et al.</i> (2007)
<i>i4g2</i>	<i>EUKARYOTIC TRANSLATION INITIATION FACTOR (ISO) 4G 2</i>			Recessive <i>r</i> gene, initiation of viral RNA translation	More resistant, missing component for viral gene expression	Nicaise <i>et al.</i> (2007)
<i>jin1</i>	<i>JASMONATE INSENSITIVE 1</i> (AT1G32640)			Loss of JA signaling; negative regulator of SA-dependent signaling	More resistant, enhanced SAR	Laurie-Berry <i>et al.</i> (2006)
<i>npr1-1</i>	<i>NONEXPRESSER OF PR GENES 1</i> (AT1G64280)			Ankyrin-repeat protein required for <i>PR-1</i> activation	More susceptible, no SAR, no ISR	Cao <i>et al.</i> (1994)
<i>p58^{IPK}</i>	<i>HOMOLOG OF MAMMALIAN P58^{IPK}</i> (AT5G03160)			Constitutive activation of PKR	More resistant, strong apoptosis-mediated HR	Bilgin <i>et al.</i> (2003)
<i>pad4-1</i>	<i>PHYTOALEXIN DEFICIENT 4</i> (AT3G52430)			Lipase-like protein, positive regulator pathogen-dependent SA signaling	More susceptible, no SAR	Cui <i>et al.</i> (2018)
<i>sid2-1</i>	<i>SA INDUCTION DEFICIENT 2</i> (AT1G74710)			Isochorismate synthase, required for SA biosynthesis	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

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.

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.

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	<i>p58^{IPK}/L4</i>	A233G	synonymous			
	<i>dcl2 dcl4/L1</i>	U289C	V45A	conservative nonpolar by nonpolar		
	<i>eds8-1/L3</i>	G298U	G75V	conservative nonpolar by nonpolar		
	<i>p58^{IPK}/L1</i>	A490G	E121G	negatively charged by nonpolar	3:5	7:1
	<i>npr1-1/L3</i>	C633U	P169S	nonpolar by nonpolar, strong structural effect		
	<i>eds8-1/L3</i>	U692C	synonymous			
	<i>i4g2/L3</i>	C806U	synonymous			
	<i>hsp90-1/L4</i>	G1087A	G320E	nonpolar by negatively charged		
HC-Pro	<i>dbp2/L4</i>	G1273A	S20N	small polar by larger polar		
	<i>cpr5-2/L4</i>	A1430G	synonymous			
	<i>dbp2/L4</i>	G1512A	D100N	negatively charged by polar		
	<i>dcl2 dcl4/L1</i>	U1520C	synonymous		5:4	9:0
	<i>jin1/L4</i>	C1760U	synonymous			
	<i>jin1/L5</i>					
	<i>npr1-1/L3</i>	C1916U	synonymous			

	<i>i4g2/L4</i>	G2316A	V367I	conservative small nonpolar by longer nonpolar		
	<i>dbp2/L2</i>	A2493G	T326S	small polar by nonpolar		
P3	<i>dbp2/L3</i>	C3205U	A206V	conservative nonpolar by nonpolar		
	<i>jin1/L4</i>	U3269C	synonymous			
	<i>jin1/L5</i>				2:3	3:1
	<i>cpr5-2/L5</i>	C3541U	T318M	small polar by long nonpolar		
	<i>cpr5-2/L5</i>	A3564U	T326S	conservative polar by polar		
6K1	<i>cpr5-2/L5</i>	U3686C	synonymous			
	<i>p58^{IPK}/L5</i>	U3743C	synonymous		2:0	2:0
C1	<i>dcl2 dcl4/L5</i>	G4032A	A72T	small nonpolar by polar		
	<i>dcl2 dcl4/L5</i>	G4148A	synonymous			
	<i>cpr5-2/L4</i>	G3269A	V148I	conservative small nonpolar by larger nonpolar		
	<i>cpr5-2/L5</i>					
	<i>npr1-1/L1</i>	A4357G	N186S	positively charged by small polar		
	<i>npr1-1/L3</i>	U4409G	I200M	conservative nonpolar by nonpolar	6:13	15:4
	<i>i4g2/L5</i>	G4448A	synonymous			
	<i>jin1/L1</i>	C4559U	synonymous			
	<i>jin1/L2</i>	A4674G	I291V	conservative nonpolar by nonpolar		
	<i>dcl2 dcl4/L1</i>	A4941G	I378V	conservative nonpolar by nonpolar		

	<i>jin1</i> /L1	C4945A	P379Q	nonpolar by polar, strong structural effect		
	<i>cpr5-2</i> /L4	G5183A	synonymous			
	<i>cpr5-2</i> /L5					
	<i>i4g2</i> /L3	G5193U	A416S	nonpolar by polar		
	<i>npr1-1</i> /L1	G5213A	synonymous			
	<i>i4g2</i> /L1	U5372C	synonymous			
	<i>i4g2</i> /L4					
	<i>i4g2</i> /L1	C5193G	A532G	conservative nonpolar by nonpolar		
	<i>p58^{IPK}</i> /L2	U5519C	synonymous			
6K2	<i>jin1</i> /L3	A5900G	synonymous		1:0	1:0
	<i>eds8-1</i> /L4	C5977U	H33Y	positively charged by polar		
	<i>eds8-1</i> /L5					
	<i>hsp90-1</i> /L1	C6219U	L107F	conservative nonpolar by nonpolar		
VPg	<i>dcl2 dcl4</i> /L4				0:47	45:2
	<i>dbp2</i> /L1					
	<i>dbp2</i> /L5	G6237A	D113N	negatively charged by polar		
	<i>i4g2</i> /L2					
	<i>i4g2</i> /L5					

<i>dcl2 dcl4/L2</i>			
<i>dcl2 dcl4/L3</i>			
<i>eds8-1/L4</i>			
<i>p58^{IPK}/L4</i>			
<i>npr1-1/L2</i>			
<i>eds8-1/L2</i>	G6237C	D113H	negatively charged by positively charge
<i>cpr5-2/L3</i>			
<i>cpr5-2/L4</i>			
<i>cpr5-2/L5</i>	A6238G	D113G	negatively charged by nonpolar
<i>npr1-1/L1</i>			
<i>npr1-1/L5</i>			
<i>hsp90-1/L2</i>			
<i>npr1-1/L4</i>	U6239A	D113E	conservative negatively charged by negatively charged
<i>i4g2/L3</i>	A6243C	N115H	polar by positively charged
<i>eds8-1/L3</i>			
<i>dbp2/L1</i>			
<i>dcl2 dcl4/L5</i>	A6243G	N115E	polar by negatively charged
<i>p58^{IPK}/L1</i>			
<i>dbp2/L2</i>	A6244G	N115S	conservative polar by polar
<i>dbp2/L3</i>			

<i>dbp2</i> /L4			
<i>jin1</i> /L3			
<i>eds8-1</i> /L1			
<i>hsp90-1</i> /L5	A6247G	E116G	negatively charged by small nonpolar
<i>p58^{IPK}</i> /L2			
<i>i4g2</i> /L1			
<i>i4g2</i> /L4	U6252C	R118C	positively charged by nonpolar
<i>dcl2 dcl4</i> /L1			
<i>p58^{IPK}</i> /L3			
<i>jin1</i> /L1			
<i>jin1</i> /L2	G6253A	R118H	conservative positively charged by positively charged
<i>jin1</i> /L4			
<i>jin1</i> /L5			
<i>hsp90-1</i> /L3			
<i>hsp90-1</i> /L4			
<i>cpr5-2</i> /L1	A6261G	K121E	positively charged by negatively charged
<i>npr1-1</i> /L3			
<i>eds8-1</i> /L4			
<i>p58^{IPK}</i> /L5			

Nla-Pro	<i>npr1-1/L4</i>	G6950A	synonymous			
	<i>dcl2 dcl4/L3</i>	A7002U	I176F	nonpolar by nonpolar	2:2	3:1
	<i>dcl2 dcl4/L3</i>	G7029A	D187N	negatively charged by polar		
	<i>jin1/L3</i>	C7181U	synonymous			
Nlb	<i>i4g2/L5</i>	A7394G	synonymous			
	<i>eds8-1/L3</i>	C8271U	synonymous			
	<i>eds8-1/L5</i>					
	<i>npr1-1/L5</i>	U8302C	L366P	nonpolar by nonpolar, strong structural effect	5:3	7:1
	<i>npr1-1/L5</i>	C8308U*	T368I	polar by nonpolar		
	<i>dbp2/L3</i>	C8337G	L378V	conservative nonpolar by nonpolar		
	<i>i4g2/L1</i>	U8345C	synonymous			
	<i>dbp2/L3</i>	G8360A	synonymous			
CP	<i>hsp90-1/L2</i>	G8792A	synonymous			
	<i>jin1/L3</i>					
	<i>jin1/L4</i>					
	<i>jin1/L5</i>	G8965A	S70N	conservative polar by polar	1:10	1:10
	<i>dcl2 dcl4/L1</i>					
	<i>npr1-1/L4</i>	A9091G	D112G	negatively charged by small nonpolar		
	<i>i4g2/L5</i>	A9091C	D112A	negatively charged by small nonpolar		

<i>cpr5-2/L4</i>	G9224A	M156I	conservative nonpolar by nonpolar	1162
<i>dcl2 dcl4/L2</i>	A9244G	E163G		
<i>dbp2/L5</i>	A9252G	I166V	negatively charged by small nonpolar	
<i>hsp90-1/L1</i>	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.