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

Rebeca Navarro,1,†,‡ Silvia Ambrós,1,† Anamarija Butković,3,4,†† José L. Carrasco,1,†† Rubén González,1,††,§§§ Fernando Martínez,1,§§§ Beilei Wu,1,‡‡‡ and Santiago F. Elena1,2,†††

1Instituto de Biología Integrativa de Sistemas (CSIC—Universitat de València), Paterna, València 46182, Spain; 2The Santa Fe Institute, Santa Fe, NM 87501, USA; 3Equal contribution.

Abstract

It is assumed that host genetic variability for susceptibility to infection conditions virus evolution. Differences in host susceptibility can drive a virus to diversify into strains that track different defense alleles (e.g. antigenic diversity) or to infect only the most susceptible genotypes. Here, we have studied how variability in host defenses determines the evolutionary fate of a plant RNA virus. We performed evolution experiments with Turnip mosaic potyvirus in Arabidopsis thaliana mutants that had disruptions in infection-response signaling pathways or in genes whose products are essential for potyvirus infection. Plant genotypes were classified into five phenogroups according to their response to infection. We found that evolution proceeded faster in more restrictive hosts than in more permissive ones. Most of the phenotypic differences shown by the ancestral virus across host genotypes were removed after evolution, suggesting the combined action of selection and chance. When all evolved viral lineages were tested in all plant genotypes used in the experiments, we found compelling evidences that the most restrictive plant genotypes selected for more generalist viruses, while more permissive genotypes selected for more specialist viruses. Sequencing the genomes of the evolved viral lineages, we found that selection targeted the multifunctional genome-linked protein VPg in most host genotypes. Overall, this work illustrates how different host defenses modulate the rates and extent of virus evolution.

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

1. Introduction

The spectrum of disease severity can be attributed to heterogeneity in virus virulence or host resistance and/or tolerance factors; the two are not necessarily independent explanations and they must actually complement and/or interact with each other resulting in an arms-race coevolutionary process. A common situation faced by viruses is that host populations consist of individuals that have different degrees of permissiveness to infection (Schmid-Hempel and Koella 1994; Pfennig 2001; Sallinen et al. 2020). Therefore, adaptive changes improving viral fitness in one host genotype may be selected against, or be neutral, in an alternative one. Genetic variability in the permissiveness of hosts and infectiousness of viruses have been well studied in animals and plants (e.g. Schmid-Hempel and Koella 1994; Altizer et al. 2006; Hughes and Boomsma 2006; Brown and Tellier 2011; Antilla et al. 2015; Parrat, Numminen, and Laine 2016; González, Butković, and Elena 2019; Sallinen et al. 2020). Theoretically, in the absence of host heterogeneity, parasites must evolve toward a host exploitation strategy that maximizes transmission with low virulence (Haraguchi and Sasaki 2000; Regoes, Nowak, and Bonhoeffer 2000; Rodriguez and Torres-Sorando 2001; Gunosov, Bergstrom, and Antia 2002; Gandon 2004; Lively 2010; Moreno-Gámez, Stephan, and Tellier 2013).
The interaction between host and parasite genotypes has been explained in the light of two different models that represent the two extremes on a continuum of possibilities. At one extreme, the gene-for-gene model, in which a parasite genotype that can infect all host genotypes exists implies that a universally susceptible host genotype should also exist (Flor 1956). Resistance occurs when a host’s resistance gene is matched by at least one parasite’s avirulence gene. Polymorphism in permissiveness and infectivity can be maintained only if virulence pays a cost. At the other extreme, the matching-alleles model is based on self- versus nonself-recognition systems. Infection is not possible unless the parasite possesses all alleles that match those of the host (Frank 1993). In this case, polymorphisms are maintained by negative frequency-dependent selection. Along the continuum, variability in host defense mechanisms impacts the evolution of the parasite’s virulence and other fitness-related traits. 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 that limit the spread of the virus and result in resistance (Zhou and Zhang 2020). Broadly speaking, these factors can be classified into basal if they are preexisting and limit within-cell propagation and cell-to-cell spread and inducible if they are only activated upon infection and inhibit systemic virus movement and replication. Basal mechanisms include susceptibility (S) genes that involve alleles of cellular proteins that do not interact properly with viral factors, e.g. translation initiation factors required by the virus for successful exploitation of the cell’s protein synthesis machinery, heat shock proteins that assist the formation of multiprotein complexes, or DNA-binding phosphatases (Carr, Lewsey, and Palukaitis 2010; Mäkinen 2019). In contrast, inducible mechanisms include genes whose expression results 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, Burch-Smith, and Dinesh-Kumar 2005; Carr, Lewsey, and Palukaitis 2010). These changes include local cell apoptosis (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, Chandra-Shekara, and Klessig 2006; Carr, Lewsey, and Palukaitis 2010), and the activation of the RNA-silencing-based resistance, that seems to play a role both in basal and inducible mechanisms (Voinnet 2001; Carr, Lewsey, and Palukaitis 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, and production of nitric oxide (Soosaar, Burch-Smith, and Dinesh-Kumar 2005). These immediate changes are followed by cell apoptosis 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 interactions between plants and beneficial microbes. Indeed, it appears that ISR is not effective against most viruses (Ton et al. 2002; Loebenstein 2009; Pieterse et al. 2009). Both SAR and ISR pathways converge into two master regulator genes, the ENHANCED DISEASE SUSCEPTIBILITY 1 (EDS1) and the PHYTOALEXIN DEFICIENT 4 (PAD4). EDS1 and PAD4 repress ISR and promote SAR. Although the SA and JA/ET pathways have been 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 with the RNA-silencing pathway (Soosaar, Burch-Smith, and Dinesh-Kumar 2005; Carr, Lewsey, and Palukaitis 2010; Yang et al. 2020). This crosstalk shows the topological structure of an incoherent feed-forward loop that confers robustness and tunability to the plant immune network (Mine et al. 2017).

Here we have used the turnip mosaic virus (TuMV; species Turnip mosaic potyvirus, genus Potyvirus, family Potyviridae)—Arabidopsis thaliana experimental pathosystem to explore how defects in basal and inducible defenses affect virus evolution. After exploring the variability of phenotypic responses to TuMV infection in a collection of twenty-one A. thaliana genotypes, ten of them were chosen, covering a wide spectrum of host phenotypic responses. Then, we performed virus evolution experiments on each of these genotypes and tracked the evolution of quantitative disease-related traits. At the end of the evolution experiment, we evaluated the effects of host genotypes in the extent of observed phenotypic changes, rates of virus evolution and the contribution of historical contingency, selection and stochasticity in the outcome of evolution. Next, we explored whether different TuMV lineages evolved as more specialist or more generalists depending on the disrupted defense mechanism in their local host. Finally, we identified the pattern of molecular changes in the genomes of the different viral lineages.

2. Materials and methods

2.1 Plants, viruses, and growth conditions

A collection of twenty-one different A. thaliana (L.) Heynh genotypes of the Col-0 accession were used for this study Table 1. This collection covered a wide range of possible resistance mechanisms. None of them showed a strong deleterious phenotype that might later interfere with the identification of infection symptoms. In all experiments described below, plants were maintained in a BSL2 climatic chamber under a photoperiod of 8-hour light (LED tubes at PAR 90–100 μmol/m²/s) at 24°C and 16-hour dark at 20°C and 40 per cent relative humidity.

TuMV infectious sap was obtained from TuMV-infected Nico-tiana benthamiana Domi plants inoculated with the infectious plasmid p35STunos containing a cDNA of the TuMV genome (GeneBank accession AF530055.2) under the control of the cauliflower mosaic virus 3SS promoter and the nos terminator (Chen et al. 2003) as described elsewhere (González, Butković, and Elena 2019; Corrêa et al. 2020). This TuMV sequence variant corresponds to the YCS strain from calla lily (Zantedeschia sp.) (Chen et al. 2003). After plants showed symptoms of infection, they were pooled and frozen with liquid N₂. This frozen plant tissue was homogenized into a fine powder using a Mixer Mill MM400 (Retsch GmbH, Haan, Germany). For inoculations, 0.1 g of powder was diluted in 1 ml inoculation buffer (50 mM phosphate buffer pH 7.0, 3 per cent PEG6000, 10 per cent Carborundum) and 5 μl of the inoculum was gently rubbed onto two leaves per plant. Plants were all inoculated when they reached growth stage 3.5 in the Boyes et al. (2001) scale. This synchronization ensures that all hosts were at the same phenological stage when inoculated.

2.2 Phenotyping infection

Five different disease-related traits were measured for each infected plant (n = 20) during 15 days postinoculation (dpi): (1) change in dry weight of the aerial part of infected plants, with a precision of 10 mg, relative to the corresponding noninfected controls (Supplementary Fig. S1A). (2) Severity of symptoms evaluated...
Table 1. Different A. thaliana genotypes used in this study. Highlighted in gray are those used for the evolution experiments.

| Genotype | Gene name | Affected pathway | Expected phenotype relative to wildtype plants | Reference |
|----------|-----------|-----------------|---------------------------------------------|------------|
| Wildtype coi1-4 | CORONATINE INSENSITIVE 1 (AT2G39940) | Repression of JA-responsive genes | No ISR, no effect on virus | Thines et al. (2007) |
|          | cpr5-2 | CONSTITUTIVE EXPRESSION OF PR GENES 5 (AT5G64930) | | | Love et al. (2007) |
|          | dhp2 | DNA-BINDING PROTEIN PHOSPHATASE 2 | Transcriptional regulation of gene expression in potyvirus-infected plants | More resistant to potyvirus infection | Castelló et al. (2011) |
|          | dcl2 | DICER-LIKE 2 (AT3G03300) | Partial loss of RNA-silencing | No effect, siRNAs produced by DCL4 | Bouché et al. (2006) |
|          | dcl4 | DICER-LIKE 4 (AT5G20320) | Partial loss of RNA-silencing | No effect, siRNA produced by DCL4 | Bouché et al. (2006) |
|          | dcl2 dcl4 | Double mutant dcl2 dcl4 | Complete loss of RNA-silencing | More susceptible, no siRNA production | Bouché et al. (2006) |
|          | dip1 | DBP-INTERACTING PROTEIN 2 (AT5G03210) | Transcriptional regulation of gene expression in potyvirus-infected plants | More susceptible | Castelló et al. (2011) |
|          | eds4-1 | ENHANCED DISEASE SUSCEPTIBILITY 4 (AT5G12000) | Loss of SA-dependent signaling | More susceptible, no SAR | Gupta, Willits, and Glazebrook (2000) |
|          | eds5-1 | ENHANCED DISEASE SUSCEPTIBILITY 5 (AT4G39030) | Lipase-like protein, positive regulator of pathogen-dependent SA signaling | More susceptible, no SAR | Nawrath et al. (2002) |
|          | eds8-1 | ENHANCED DISEASE SUSCEPTIBILITY 8 | Reduced expression of plant defense genes, reduced ISR | More resistant, enhanced SAR | Love et al. (2007) |
|          | ein2-1 | ETHYLENE INSENSITIVE 2 (AT5G05280) | MAPK, ET signaling intermediate, negative regulator SA-dependent signaling | More resistant, enhanced SAR | Love et al. (2007) |
|          | etr1-1 | ETHYLENE RESPONSE 1 (AT1G66340) | ET receptor, negative regulator SA-dependent signaling | More resistant, enhanced SAR | Love et al. (2007) |
|          | hsp90-1 | HEAT SHOCK PROTEIN 1 | Recessive gene, required for membrane-bound replication complexes, protein folding | More resistant, missing component for viral replication | Verchot (2012) |
|          | i4g1 | EUKARYOTIC TRANSLATION INITIATION FACTOR (ISO) 4 G 1 (AT3G60240) | Recessive gene, initiation of viral RNA translation | More resistant, missing component for viral gene expression | Nicaise et al. (2007) |
|          | i4g2 | EUKARYOTIC TRANSLATION INITIATION FACTOR (ISO) 4 G 2 | Recessive gene, initiation of viral RNA translation | More resistant, missing component for viral gene expression | Nicaise et al. (2007) |
|          | jin1 | JASMONATE INSENSITIVE 1 (AT1G32640) | Loss of JA signaling, negative regulator of SA-dependent signaling | More resistant, no SAR, no ISR | Laurie-Berry et al. (2006) |
|          | npr1-1 | NONEXPRESSER OF PR GENES 1 (AT1G64280) | Ankyrin-repeat protein required for PR-1 activation | More resistant, strong apoptosis-mediated HR | Bilgin et al. (2003) |
|          | p58PK | HOMOLOG OF MAMMALIAN P58IPX (AT5G03160) | Constitutive activation of PKR | More susceptible, no SAR | Cui et al. (2018) |
|          | pad4-1 | PHYTOALEXIN DEFICIENT 4 (AT5G02430) | Lipase-like protein, positive regulator pathogen-dependent SA signaling | More susceptible, no SAR | Nawrath and Métraux (1999) |
|          | sid2-1 | SA INDUCTION DEFICIENT 2 (AT1G74170) | Isochorismate synthase, required for SA biosynthesis | More susceptible, no SAR | |
addition, two lineages were also evolved in wild type (WT) plants. To begin the evolution experiment, ten *A. thaliana* plants at developmental stage 3.5 in Boyes et al. (2001) scale per lineage and genotype were inoculated as described above with the virus stock previously generated from infected *N. benthamiana* plants inoculated with infectious clone p35STunos. The next passages were made by harvesting the symptomatic plants at 12 dpi, preparing the infectious sap as described above, and inoculating it to a new healthy population of ten plants.

2.4 Total RNA extractions and quantification of viral load

Pools were made of ten infected symptomatic plants per lineage, genotype, and serial passage, frozen with liquid N₂ and preserved at −80°C until it was homogenized into fine powder. Next, aliquots of approximately 0.1 g of ground tissues were used for total RNA (RNAi) extractions. The RNAi was extracted with the Agilent Plant RNA isolation Mini kit (Agilent Technologies, Santa Clara CA, USA). Three aliquots of RNAi per each sample were separated and their concentration adjusted at approximately 50 ng/µl to estimate viral load.

Viral load was quantified in each of the three aliquots by absolute real-time quantitative RT-PCR (RT-qPCR) using standard curves and the primers TuMV F 117 forward (5′-CGAGTGATCTTATAGCCACGCTGGACACC′-3′) and F 118 reverse (5′-TAACCCCCCTTAAACGCCCA GTAAAG′-3′) that amplify a 173 nucleotides fragment from the 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 RNAi 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 in an ABI StepOne Plus Real-time PCR System (Applied Biosystems, Foster City CA, USA). The cycling conditions consisted of an RT phase of 5 min at 42°C and 10 min at 95°C followed by a PCR stage consisting of 40 cycles of 5 s at 95°C and 34 s at 60°C, and the final melt curve profile that consisted of 15 s at 95°C, 1 min at 60°C and 15 s at 95°C. Negative controls consisted of healthy RNAi plant extract (mock-inoculated noninfected control) and water. For each sample, three technical replicates were quantified. The results were analyzed using the StepOne software 2.2.2 (Applied Biosystems).

2.5 TuMV genome amplifications and sequencing

A sample of the above RNAi obtained at passage 12 from each viral lineage was used to generate the viral consensus sequence. RNAs were amplified by high-fidelity RT-PCR 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 RNAi (150–300 ng) was mixed with 0.25 µM of the 1 R-P3 (5′-CGAGTGATCTTATAGCCACGCTGGACACC′-3′), 2 R-Nla (5′-TGCTCGAATAGCGAATGTAGCGATTGGTG-3′) or 3 R-polyAR (5′-GTTTTTTTTTTTTTTTTTTGCCTGGACACC′-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 a mix containing AccuScript Hi-Fi 1× Buffer, 1 mM of dNTPs, 8 mM of DTT, 4 U of Ribolock RNase inhibitor (Thermo Scientific), and 0.5 µl of AccuScript Hi-Fi (Agilent Technologies) was added up to a 10 µl volume. RT reactions consisted 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 µM of DMSO, 2 U of Phusion DNA polymerase (Thermo Scientific), and 1 µl of the corresponding RT reaction. R1 fragment was amplified using the primer set 1F-SUTR (5′-GCAAAAGCCAGCATTTGGCGACACTCAAGC-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′-TGCGGAGCTTTTGCGGATGGTAGATACACATT-3′) and 2 R-Nla or 3R-Nla (5′-CTCGTTATATAGGAGCTGGATATACACATT-3′) and 3R-polyAR, respectively and a PCR with the same denaturation and extension steps than the fragment R1 but different amplification steps: a stage consisting of 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. For fragment R3, 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 were used. PCR products were purified with the MSB Spin PCRampe Kit (Strategic 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.

2.6 Statistical analyses

AUDPS data were fitted to an analysis of variance (ANOVA) model with plant genotype (G) as main factor, 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:

$$ \text{AUDPS}_{ijk} (t) \sim A + t + G_i + L(G)_{ij} + (t \times G)_i + (t \times L(G))_{ij} + \varepsilon_{ijk} \quad (1) $$
where $\text{AUDPS}_{ik}(t)$ is the value observed at time $t$, for an individual infected plant $k$ of evolutionary lineage $j$ of genotype $i$. $A$ represents the grand mean value and $\epsilon_{ik}$ stands for the error assumed to be Gaussian distributed at every $t$. The type III sum of squares was used to partition total variance among factors. The magnitude of the effects was evaluated using the $\eta_j^2$ statistic (proportion of total variability in AUDPS attributable to each factor in the model; conventionally, values of $\eta_j^2 \geq 0.15$ are considered as large effects).

To test the relative contribution of selection, chance, and historical contingency in the outcome of the evolutionary process, the statistical approach described by Travissano et al. (1995) has been used. One can imagine two situations. Firstly, ancestral biological differences among phenotypes could be preserved during 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 $\approx 1$, the greater the importance of ancestral differences. Secondly, 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 $< 1$. The lesser the effect of ancestral differences, the flatter the slope, being zero in the extreme case where ancestral differences have been completely erased. The estimated slope of the regression of the evolved phenotypic values against the ancestral ones was obtained by least-squares linear regression. The estimated slopes were compared against the slope expected under the null hypothesis of equal size with bounds

$$
\left[ \frac{\text{AUDPS}_{\text{min}} + \frac{i}{10} (\text{AUDPS}_{\text{max}} - \text{AUDPS}_{\text{min}})}{\text{AUDPS}_{\text{min}}} \right] \leq \eta_i \leq \left[ \frac{\text{AUDPS}_{\text{max}} - \text{AUDPS}_{\text{min}}}{\text{AUDPS}_{\text{max}} - \text{AUDPS}_{\text{min}}} \right]
$$

where $i \in [0, 9]$, as described in Moury et al. (2021).

### 2.8 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, $\pi_i$, the nucleotide diversity for the entire sample, $\pi$, the interhost genotypes nucleotide diversity, $\delta_i$, and the estimate of the proportion of interhost genotypes nucleotide diversity, known as coefficient of nucleotide differentiation (Nei 1982), $N_{ST} = \delta_i / \pi_i$. Standard deviations of estimates were inferred from 1,000 bootstrap samples. All these computations were done using MEGA 11 (Tamura, Stecher, and Kumar 2021) and the lowest-BIC nucleotide substitution model Kimura 2-parameters (Kimura 1980). Tajima’s D test of selection (Tajima 1989) and its statistical significance were evaluated using DnaSP6 (Rozas et al. 2017).

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

$$
\text{probit}(m_i) = \mu + C_i + \epsilon_i
$$

where $\mu$ is the average genomic mutation frequency and $i$ refers to the ten cistrons in the main open reading frame.

### 3. Results

#### 3.1 Classification of plant genotypes according to the phenotype of TuMV infection

The twenty-one $A. \text{thaliana}$ genotypes used in this 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 WT plants. Genotypes were classified according to the phenotype of TuMV infection based on the five different disease-related traits described in Section 2.2. In order to classify plant genotypes according to their similarity in phenotypic response to TuMV infection, a nearest-neighbor clustering was obtained Fig. 1. We found five significant phenogroups that also minimized the BIC in a $k$-means analysis [BIC(5) = 12.241 compared to BIC(4) = 15.668 and BIC(6) = 18.249]. Hereafter, we will refer to them as phenogroups G1 to G5. The characterization of the phenotypic differences among these groups is presented in the Supplementary Text S1 and Figs. S1 and S2. In short, the
members of phenogroup G1 show enhanced restriction to TuMV infection. The only member of phenogroup G2 (pS8GR) shows no significant changes in AUDPS, infectivity, symptoms severity, and viral load relative to infected WT plants but a strong reduction in the change in dry weight. The only member of phenogroup G3 (dbp2) also shows enhanced restriction although without significant reduction in viral load. G4 represents a sort of hodgepodge formed by genotypes that do not show a clear phenotypic difference from WT. The only member of phenogroup G5 (jin1) shows a significant increase in all measured traits except viral load.

Representatives from the five phenogroups were randomly selected for the subsequent evolution experiment (shadowed rows in Table 1): eds8-1 and J4g2 from phenogroup G1 representing the more restrictive genotypes; pS8GR from G2; dbp2 from G3; cpr5-2, dcl2 dcl4, hasp90-1, npr1-1, and WT from G4; and jin1 from G5 representing the more permissive genotype.

3.2 Experimental TuMV evolution in plants from each phenogroup

First, we sought to evaluate the dynamics of evolution for AUDPS Fig. 2. Recall that in our experiments, only viruses evolved while plants did not, so changes in infection phenotypes were due to the evolution of viruses and the way they interact with plants. Notice that lineage cpr5-2/L2 showed quite a different evolution pattern from the other four lineages evolved in cpr5-2. It never increased in AUDPS and was lost after passage five Fig. 2. Therefore, we ended up with a total of forty-six evolved lineages.

Data were fitted into the ANOVA model described in Eq. 1 of Section 2.6. Despite a considerable amount of noise in the time series, we found some significant results that can be summarized as follows Table 2. First, a net significant and large $\eta^2_T = 0.399$ effect associated with the evolutionary passages (t term in Eq. 1) has been observed, indicating that AUDPS evolved during the experiment. Second, plant genotypes (G term in Eq. 1) had a highly significant and very large $\eta^2_G = 0.868$ effect on the phenotypic traits, suggesting that TuMV evolutionary dynamics were strongly influenced by the local host genotype in which it was being passed. Third, independent lineages evolved in the same host genotype show a high degree of parallelism, as indicated by a nonsignificant lineage effect [t(G) term in Eq. 1, with $\eta^2_L = 0.029$; and term t x L(G) in Eq. 1, with $\eta^2_L = 0.039$].

To evaluate the role of the ancestral differences (Travisano et al. 1995) in the evolution of AUDPS across the ten plant genotypes, we evaluated the magnitude of their change at the end of the evolution experiment. Fig. 3A shows the plots of evolved vs. ancestral values. For illustrative purposes, the solid red line represents the null hypothesis of absolute preservation of ancestral differences. A significant regression exists between evolved and ancestral values ($R = 0.491$, $F_{4,44} = 14,000$, $P < 0.001$), with slope $0.284 \pm 0.076$ ($\pm 1$ SD). Since the slope is still significantly different from zero, yet clearly flatter than the expected relationship of slope one ($4.422 , P < 0.001$), we concluded that ancestral differences have been mostly removed by the combined action of selection and chance, yet not completely erased.

3.3 Rates of AUDPS evolution

To explore the possible impact of different plant immunity mechanisms on the rates of virus phenotypic evolution, we estimated the rates of AUDPS evolution using the power-law model described by Eq. 2 in Section 2.6.

Fig. 3B shows the estimated rates of evolution for host genotypes ordered from the most restrictive (phenogroup G1) to the most permissive (phenogroup G5). Very interestingly, the regression coefficient was significantly negative ($-6.394 \cdot 10^{-4}$ $\pm 2.843 \cdot 10^{-4}$; $R = 0.321$, $F_{4,44} = 5.061$, $P = 0.030$), suggesting that evolution proceeded faster in the most restrictive hosts and slower in the most permissive ones.

In conclusion from this section, we found that differences in the maximum rates of evolution at the beginning of the evolution experiment between permissive and resistant hosts actually depend on the particular resistance mechanism.
Table 2. Results of the ANOVA model for AUPDS evaluated along the course of experimental evolution.

| Source of variation | Type III Sum of squares | d.f. | Mean squares | F     | P     | η² | 1—β |
|---------------------|------------------------|------|--------------|-------|-------|----|-----|
| A                   | Hypothesis             | 5.960| 1            | 5.960 | 1255.393 | < 0.001 | 0.962 | 1   |
| Error               |                        | 0.236|              | 49.783| 0.005  |     |     |     |
| t                   | Hypothesis             | 3.705| 1            | 3.705 | 337.122 | < 0.001 | 0.399 | 1   |
| Error               |                        | 0.11 |              | 508   | 0.011  |     |     |     |
| G                   | Hypothesis             | 1.299| 9            | 0.144 | 31.668 | < 0.001 | 0.868 | 1   |
| Error               |                        | 0.005|              | 43.293| 0.005  |     |     |     |
| t × G               | Hypothesis             | 0.144| 9            | 0.016 | 1.458  | 0.161  | 0.025 | 0.700|
| Error               |                        | 0.011|              | 508   | 0.011  |     |     |     |
| L(G)                | Hypothesis             | 0.167| 38           | 0.004 | 0.399  | 1     | 0.029 | 0.439|
| Error               |                        | 0.011|              | 508   | 0.011  |     |     |     |
| t × L(G)            | Hypothesis             | 0.224| 37           | 0.006 | 0.550  | 0.986  | 0.039 | 0.609|
| Error               |                        | 0.011|              | 508   | 0.011  |     |     |     |

η²: magnitude of the effect; 1—β: power of the test.
The different factors are defined in equation 1: t: evolutionary passage; G: plant genotype; L: virus evolutionary lineage; t × G interaction between evolutionary passage and plant genotype; L(G) viral evolving lineage within-host genotype, and t × L(G) interaction between evolutionary passage and viral evolving lineage.

3.4 The degree of permissiveness of plant genotypes to infection drives the evolution of TuMV lineages into more or less specialization

To analyze the specificity of adaptation of each evolved TuMV lineage, we performed a full cross-infection experiment. In this experiment, the forty-four viral lineages evolved in the mutant genotypes (excluding the two WT-evolved lineages) were inoculated to ten plants of the nine A. thaliana mutant genotypes used in the evolution experiments (excluding WT plants). The presence of symptoms in each plant was recorded daily for up to 12 dpi. This dataset was 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 of nestedness, modularity and specialization of a bipartite infection network (Weitz et al. 2013). Both approaches drive to identical conclusions. For the sake of keeping the presentation concise, the results of the first approach are provided in Supplementary Text S2 and Table S1.

For the second analytical approach, we used an infection matrix of forty-four rows (viral lineages evolved in each local host genotype) by nine columns (test genotypes). Fig. 4 shows the packed infection matrix with rows and columns organized to better highlight its nestedness and modularity (Weitz et al. 2013). The infection network was significantly nested (temperature nestedness: T = 1.142, P < 0.001), with genotypes jin1 and dcl2 dcl4 being permissive to most viral lineages (universally susceptible hosts), while the most restrictive genotypes eds8-1 and p58IPK were more efficiently infected by lineages that were precisely evolved in these two genotypes. Likewise, the most generalist viral lineages were eds8-1/L4, p58IPK/L3 and p58IPK/L4 that infected the nine mutant host genotypes Fig. 4 with high efficiency (universally virulent viruses). On another hand, the most specialized viral lineages that were able to more efficiently infect only their local hosts, were all surviving lineages evolved in cpr5-2, four lineages evolved in dcl2 dcl4 and four lineages evolved in jin1 Fig. 4. Therefore, we conclude that more restrictive host genotypes tended to select for
more generalist viruses while more permissive host genotypes did so for more specialized viruses, in agreement with a gene-for-gene infection model (Weitz et al. 2013).

Next, we computed Blüthgen, Menzel, and Blüthgen (2006) specialization indexes for the packed matrix Fig. 4. Firstly, we evaluated the standardized species-level measure of partner diversity $d'$. $d'$ (last column in Fig. 4) ranged from $<0.01$ for the more generalist evolved lineages (eds8-1/L4, and p58PK/L3) to $>0.1$ for the more specialist ones (dcl2 dcl4/L1 and jin1/L1–L4). Very interestingly, $d'$ values show a strongly significant negative correlation with the rank order degree of permissiveness of each host genotype, i.e. the order in which hosts appear in the columns of Fig. 4.

Figure 4. Analysis of the full cross-infection matrix. Packed matrix that highlights its nested structure, compatible with a gene-for-gene infection model. Last column shows the species-level specialization index ($d'$ = 0 for most generalists and $d'$ = 1 for most specialists). Evolved TuMV lineages are colored according to the phenogroup of their local host genotype according to the color palette used in Fig. 1.
conclude that minor yet significant genetic differentiation has significantly greater than zero ($z = 5.916$, $H'_s = 3.835$, $H'_m = 5.916$, $P < 0.001$), meaning that, overall, the binary infection network shows a degree of specialization. Most of the lineages were able of infect one or a few host genotypes, a result mainly driven by those lineages evolved in the less restrictive host genotypes.

Finally, we sought to explore the modularity of the infection network. In this context, a module will refer to an aggregated set of viral lineages and their hosts characterized by more interactions within the module than between modules (Newman 2006; Dormann and Strauss 2014), as expected for a mutation accumulation infection model (Weitz et al. 2013). We computed the Q modularity index, that 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 ($Q = 0.066$, $P < 0.001$) in the infection matrix Fig. 4: genotypes with mutations affecting the same signaling pathway may be selecting for viral lineages with similar properties, thus being able to infect a subset of plant genotypes with equal efficiency. An alternative hypothesis may be parallel evolution of lineages evolved in the same host genotype; the most representative cases of this possibility are the lineages evolved in cpr5-2, dcl2 dcl4 and jin1. To distinguish between these two hypotheses, we computed a reduced matrix averaging the observed AUDPS among lineages evolved in the same host genotype and transforming this nine-by-nine infection matrix into its corresponding packed binary matrix, as described above. If the reduced matrix still shows significant modularity, the first hypothesis will hold on, but if modularity vanishes then the alternative hypothesis will be more parsimonious. The reduced matrix was no longer modular ($Q = 0.029$, $P = 0.078$). This supports the notion that the observed modularity was driven by convergent phenotypic evolution of lineages evolved into the same host genotype rather than by common selective pressures owed to mutations in overlapping defense mechanism.

3.5 Genomic changes in evolved TuMV lineages

Finally, we have explored the genomic changes experienced by the TuMV evolved lineages. A total of 119 mutational events have been observed, affecting 73 nucleotide positions (Supplementary Fig. S3 and Table S2). No single mutation was shared between all five independent lineages propagated on a plant genotype and very few mutations were observed in more than two lineages. According to the type of nucleotide substitution, ninety-eight were transitions and twenty-one transversions. Regarding their effect on the protein sequence, thirty were synonymous and eighty-nine nonsynonymous. Interestingly, some mutations, including thirteen nonsynonymous and four synonymous, have been observed multiple times in independent lineages. As discussed in the Section 2.8, treating each lineage as an observation and each host genotype as a subpopulation, the average nucleotide diversity within-host genotypes, referred only to the seventy-three polymorphic sites, was $\pi_S = 0.067 \pm 0.007$. On the other hand, the nucleotide diversity for the entire sample was $\pi_T = 0.073 \pm 0.008$. Hence, the estimated interhost genotype nucleotide diversity was $\delta_T = 0.006 \pm 0.003$ and the coefficient of nucleotide differentiation (Nei 1982) $N_{ST} = 0.081 \pm 0.039$, a value significantly greater than zero ($z = 2.077$, $P = 0.019$). 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 virus genetic differentiation among plant genotypes, we performed a Tajima’s $D$ test (Tajima 1989) and found that it was significantly negative ($D = -2.496$, $P = 0.006$). This result suggests the presence of selective sweeps acting on virus populations and presence of selection.

Next, we sought to characterize the distribution of mutational events along the nine nonoverlapping cistrons Fig. 5. The frequency of mutations per cistron, relative to the length of the corresponding cistron, was fitted to the logistic regression model shown in Eq. 4 in the Section 2.8. Highly significant differences exist ($\chi^2 = 143.206$, 9 d.f., $P < 0.001$) yet entirely due to the ~11-fold larger mutation frequency observed in the VPg cistron relative to the rest of the genome Fig. 5. 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 (Supplementary Fig. S3 and Table S2).

Convergent nonsynonymous mutations are, a priori, good candidates for adaptive mutations. Mutation CP/V148I appears in two lineages evolved in cpr5-2 (Supplementary Fig. S3 and Table S2), mutation CP/S70N appear in lineage dcl2 dcl4/L1 and three lineages evolved in jin1 (Supplementary Fig. S3 and Table S2), and two different mutations affecting CP; CP/D112G appears in lineage npr1-1/L4 and CP/D112A that appears in lineage i4g2/L5, that result in a similar replacement of side chains. In addition, ten nonsynonymous mutations in VPg are shared by several lineages (Supplementary Fig. S3 and Table S2). Out of these ten cases, three seem to be particularly promising candidates and are located in a narrow region of VPg (residues 113, 115, and 118) (Supplementary Fig. S3 and Table S1). Mutations G6237A, G6237C, A6238G, and U6239A all affect the same codon, resulting in amino acid replacements VPg/D113N, VPg/D113H, VPg/D113G, and VPg/D113E, respectively.

4. Discussion

4.1 Interaction between TuMV and different plant defenses

In this study we have explored the effect that mutations in different host disease signaling pathways or recessive resistance $S$ genes have in the outcome of virus evolution. Our experimental pathosystem consisted of two well-studied components: the
model plant A. thaliana and TuMV, a virus with high prevalence in natural populations on this host (Pagán et al. 2010). Among the disease signaling pathways, we have studied genes involved in the SA, JA/ET pathways, and RNA-silencing, while 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 compared 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, the most unexpected one being the result for jin1. 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 permissive genotype to TuMV infection, showing enhanced symptoms of disease (phenogroup G5). This may reflect the fact that the function of genes involved in disease signaling 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. coi1-4) behaved as WT 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 restrictive genotypes to infection (phenogroup 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. Replication of viruses such as alfalfa mosaic virus, potato virus X, or turnip vein clearing virus was inhibited by treating plants with SA, while cucumber mosaic virus appeared 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 be highly permissive to TuMV infection. However, infected plants were, overall, hardly distinguishable from the WT plants in their response to infection (Supplementary Figs. S2 and S3). A possible explanation would be the strong suppressor activity of HC-Pro in the WT plants that effectively counteracts the defense mechanisms (Kaschau and Carrington 2003). Indeed, it has been shown that the outcome of the interplay between plant RNA-silencing response and potyviruses is mostly driven by the efficiency of the viral HC-Pro suppression activity (Lin 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 4ag2 belonging to phenogroup G1, that is defective for the elF4isoE factor, was highly restrictive to TuMV infection (Nicaise et al. 2007; Charron et al. 2008). Likewise, in agreement with previous descriptions, mutant dhp2 belonging to phenogroup G3, was restrictive to potyviruses (Castelló et al. 2011). Finally, mutant hsp90-1 showed a response to TuMV infection equivalent to 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 of the knocked down expression of gene P58PK on TuMV replication. In mammals P58PK, a tetrapeptide-repeat 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, thus favoring viral spread (Goodman et al. 2011). In our experiments, p58PKH mutant plants showed enhanced tolerance to infection by showing weaker symptoms than WT plants despite achieving the same level of virus accumulation (Supplementary Fig. S1). In sharp contrast, Bilgin et al. (2003) found that infection of pS8PK-silenced N. benthamiana and A. thaliana plants infected with tobacco mosaic virus led 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 of this gene.

An important conclusion of our study, is that the generally accepted model of plant defense signaling pathways, mainly based on experiments done with bacteria and fungi, may not be most suitable to describe interaction between plants and viruses well. By contrast, recessive S gene-based resistance seems to better explain differences in susceptibility to infection and viral accumulation, thus being a more promising target for future development of resistant plants.

4.2. Evolution of more or less specialized viral strategies depends on the host genotype

The gene-for-gene and mutation accumulation models of host—virus interaction modes represent the two ends of a continuum of possible outcomes (Agrawal and Lively 2002). The most relevant difference between both models regards the expected genetic heterogeneity in both host and virus populations. With a pure gene-for-gene interaction, the susceptible host types are expected to disappear and the resistant types are expected to dominate the population. Vice versa, the most virulent virus allele would become fixed in the population at the cost of 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, Kniskern, and Bergelson 2007)] and high virulence usually comes with a cost in terms of pathogen’s transmission (Acevedo et al. 2019). Hence, a pure gene-for-gene strategy seems unlikely to be achieved. By contrast, with a pure mutation accumulation 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, Alonso-Blanco, and García-Arenal 2008) and for immunity-related genes (Todesco et al. 2010; Van de Weyer et al. 2019; Butković et al. 2021), thus suggesting that an arms race between pathogens and plants is still ongoing but it is unclear whether it may result in pure gene-for-gene or mutation accumulation interactions or lie somewhere in between. Evolution experiments with such different pathosystems, such as TuMV-A. thaliana (González, Butković, and Elena 2019), Octosporea bayeri-Daphnia magna (Altermatt and Ebert 2008), and Serratia marcescens-Caenorhabditis elegans (Gibson et al. 2020; White et al. 2020) have produced congruent results: 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, Hillung, and Elena 2015).
Our results, as well as those by Hillung et al. (2014) and González, Butković, and Elena (2019) have shown the evolution of significantly nested bipartite 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 restrictive hosts selected for more generalist viruses, here matching the predictions of the gene-for-gene 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.

4.3 Ancestral differences in TuMV performance across host genotypes have been almost completely erased after experimental evolution

A relevant question in evolutionary biology is the extent in which ancestral differences determine the fate of evolution. Put in other words, what are the relative contributions of adaptation, chance and historical contingency to the evolution of organismal fitness (Travisano et al. 1995). Two different situations can result. Firstly, ancestral differences among phenotypes are preserved during 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). Therefore, this situation would result in a nonzero slope in a regression of the evolved phenotypic values against the ancestral ones. The closer the slope is to one, the greater the importance of ancestral differences. Secondly, if ancestral trait variation was erased from the evolved populations because the combined effect of adaptation and chance, a regression slope smaller than one would be expected. The less effect of ancestral differences, the flatter the slope, being zero in the extreme case where ancestral differences have completely vanished. In our study, we found that ancestral differences in AUDPS were not entirely removed by the combined action of selection and chance. Some minor differences still are observable for lineages evolved in the most restrictive hosts p58×pK and eds8-1 Fig. 3A.

4.4 Role of natural selection

We found evidence 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 negative D value was significant, that 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 they may rise 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: fifty-six out of the seventy-three 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 exponential growth of viral populations within individual hosts.

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 of these nonsynonymous mutations evolved in the same host genotype but others arise in different host genotypes (Supplementary Fig. S3 and Table S2). Yet without a clear association with the particular signaling pathway, mutated gene Table 1 or the phenogroup they belong to Fig. 1. Interestingly, most of these convergent mutations happened in the VPg cistron, that also turns out to be the most variable one Fig. 5. 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)4G1, and interacts with all other viral proteins (Bosque et al. 2014) and some of the host cell proteins (Martinez et al. 2016, 2020). Indeed, in previous evolution experiments with potyviruses, VPg has also been shown to be an important target of selection. For example, Aguadelo-Romero et al. (2008) found that a single amino acid replacement in VPg was enough to largely increase tobacco etch virus infectivity, severity of symptoms and viral load 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, although yeast-two hybrid assays showed that none of these mutations affected the binding of VPg with eIF(iso)4E (Gallois et al. 2010). As a final example, one of the most extensively used resistance genes against potato virus Y in commercial pepper cultivars is pur, which has many different alleles (Nicaise et al. 2007; Charron et al. 2008). The pur locus encodes for the eIF4E factor that, 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; Mourgue et al. 2004; Ayme et al. 2006).

One of the two mutations identified by Gallois et al. (2010), VPg/E116Q, affects the same protein domain as mutations VpG/D113G and VpG/R118H identified here. We found VpG/D113G in several lineages evolved in different host genotypes, while VpG/R118H was only found in the jin1 lineage. The mechanism by which these two mutations may confer a selective advantage to TuMV lineages cannot be inferred from our studies and will be the subject of a follow up paper.

4.5 Comparing these results with other pathosystems

Someone may wonder whether results reported here might reflect universal patterns of virus evolution on hosts with variable susceptibility to infection or just a particularity of the way plants deal with infections. Mongelli et al. (2022) performed very similar evolution experiments with drosophila C virus (DCV), also a picornavirus, infecting eight Drosophila melanogaster genotypes that carried mutations affecting innate immunity antiviral defense pathways. Firstly, Mongelli et al. (2022) have also observed that the extent and evolutionary dynamics of viral load and infectivity depended on the particular host genotype, with a clear tendency to increase viral load in immunity-deficient flies. Secondly, also in agreement with our observations, DCV results also highlight pleiotropy (viral genotype-by-host genotype interaction)
as a major determinant of viral evolution. Indeed, they clearly show that the fitness effect of mutations that already existed in the standing genetic variation of the stock DCV population was strongly dependent on both, the fly’s genetic background (the same mutation being beneficial or deleterious in different backgrounds) and the virus’ genetic background (epistatic interactions with other mutations in the same haplotype). Thirdly, our observation that ancestral phenotypic differences in virulence across host genotypes can be erased after a few passages in each host genotype are also in agreement with Mongelli et al. (2022) observations, which suggest that the more restrictive a host genotype is to infection the more changes of selection and evolution of generalist viruses.

To conclude, besides obvious differences among experimental systems and other subtleties, these two studies demonstrate that experimental evolution-based approaches can provide insightful information about the constraints imposed by different defense signaling pathways on virus evolution and remark the prevalence on some pathways over others in controlling not only virus replication and accumulation but also, more broadly interesting, virus evolution.

4.6 Concluding remarks
Since our evolution experiments all started with an A. thaliana naive strain of TuMV, someone may consider as a potential weakness of our study the difficulty to disentangle between general adaptations to the plant host from specific adaptations to the different mutant genotypes. Although obviously some common factors are shared by all ten plant genotypes, they differ in a well-defined component (the mutated gene) that we have shown affects their permissiveness to infection. The fact that we have found significant differences in the outcome of the evolution experiments performed in different plant genotypes, by itself, is evidence that each host genotype represents a somehow different selective environment. Unfortunately, at this stage, it is not possible to assign an adaptive value to every mutation observed in the different lineages. Doing so would require to generate them individually and in all relevant combinations in the infectious clone p55TuNos and test their effect on AUDPS across the ten host genotypes. Although doable, it is a tremendous task beyond the scope of this study. The best assessment we can do so far about adaptive values is the commonly assumed principle that convergent mutations rising in independent lineages evolved in the same host genotype must provide a host-specific fitness advantage.

Pathogens evolve to escape from host defense mechanisms. Understanding how different defense mechanisms constrain the pathogen evolution is key to understand the evolutionary dynamics of infectious diseases. Here we aimed to determine the impact of defects in host defense mechanisms on virus evolution. We observed that virus interaction with defective hosts resulted in different intensity of disease-related traits, intensity that was determined by the disrupted defensive mechanism. Virus evolution on hosts with different degree of permissiveness shows that: (1) viral rates of evolution depend on the specific defective defense mechanism of the host. (2) Viruses evolve as more generalists when evolving in hosts with stronger defenses. On the contrary, evolving in more permissive hosts results in more specialized viruses. (3) Evolution results in a reduction of ancestral genetic variation, which occurs independently of the host permissiveness. (4) Regardless of their host’s defenses, evolved viruses showed convergent mutations on the same region of their genome. Altogether, this work describes the constrains that the degree of host immunity can impose on virus evolution.

Data availability
All raw data are available at https://git.csic.es/SFElena/tumv-evolution-in-a.-thaliana-genotypes-with-deficient-immune-responses.

Supplementary data
Supplementary data is available at Virus Evolution online.

Acknowledgements
We thank Francisca de la Iglesia and Paula Agudo for excellent technical assistance and the rest of the EvolSysVir lab members for fruitful discussions.

Funding
This work was supported by grants BFU2015-65037-P and PID2019-103998GB-I00 (Agencia Estatal de Investigación - FEDER) and PROMETEU2019/012 (Generalitat Valenciana) to S.E. and by Natural Science Foundation China grant 31872414 to B.W.

Conflict of interest: None declared.

References
Acevedo, M. A. et al. (2019) ‘Virulence-driven Trade-offs in Disease Transmission: A Meta-analysis’, Evolution, 73: 636–47.
Agrawal, A., and Lively, C. M. (2002) ‘Infection Genetics: Gene-for-gene versus Matching-alleles Models and All Points in Between’, Evolutionary Ecology Research, 4: 79–90.
Agudelo-Romero, P. et al. (2008) ‘Virus Adaptation by Manipulation of Host’s Gene Expression’, PLoS ONE, 3: e2397.
Aftermann, F., and Ebert, D. (2008) ‘Genetic Diversity of Daphnia Magna Populations Enhances Resistance to Parasites’, Ecology Letters, 11: 918–28.
Altizer, S. et al. (2006) ‘Seasonality and the Dynamics of Infectious Diseases’, Ecology Letters, 9: 467–84.
Alves, K. S., and Del Ponte, E. M. (2021) ‘Analysis and Simulation of Plant Disease Progress Curves in R. Introducing the Epifitter Package’, Phytopathology Research, 3: 22.
Anttila, J. et al. (2015) ‘Environmental Variation Generates Environmental Opportunistic Pathogen Outbreaks’, PLoS ONE, 10: e0145511.
Ayme, V. et al. (2006) ‘Different Mutations in the Genome-linked Protein VPg of Potato Virus Y Confer Virulence on the Pur2 Resistance in Pepper’, Molecular Plant-Microbe Interactions, 19: 557–63.
Bascompte, J. et al. (2003) ‘The Nested Assembly of Plant-animal Mutualistic Networks’, Proceedings of the National Academy of Sciences of the USA, 100: 9389–9387.
Bedhomme, S., Hillung, J., and Elena, S. F. (2015) ‘Emerging Viruses: Why are Not Jacks of All Trades?’, Current Opinion in Virology, 10: 1–6.
Bilgin, D. D. et al. (2003) ‘FSBTPK, a Plant Ortholog of Double-stranded RNA-dependent Protein Kinase PKR Inhibitor, Functions in Viral Pathogenesis’, Developmental Cell, 4: 651–61.
Blüthgen, N., Menzel, F., and Blüthgen, N. (2006) ‘Measuring specialization in species interaction networks’, BMC Ecology, 6: 9.
Bosque, G. et al. (2014) ‘Topology Analysis and Visualization of Potyivirus Protein-protein Interaction Network’, BMC Systems Biology, 8: 129.
Bouché, N. et al. (2006) ‘An Antagonistic Function for Arabidopsis DCL2 in Development and a New Function for DCL4 in Generating Viral siRNAs’, EMBO Journal, 25: 3347–56.
Maintenance in Host-parasite Interactions’, Plant Pathology, 62: 133–41.
Moury, B. et al. (2021) ‘The Quasi-universality of Nestedness in the Structure of Quantitative Plant-parasite Interactions’, Peer Community Journal, 1: e44.
——— et al. (2004) ‘Mutations in Potato Virus Y Genome-linked Protein Determine Virulence Towards Recessive Resistances in Capsicum Annuum and Lycopersicon Hirsutum’, Molecular Plant-Microbe Interaction, 17: 322–9.
Nawrath, C. et al. (2002) ‘EDSS, an Essential Component of Salicylic Acid-dependent Signaling for Disease Resistance in Arabidopsis, Is a Member of the MATE ‘Transporter Family’, The Plant Cell, 14: 275–86.
Nawrath, C., and Métraux, J. P. (1999) ‘Salicylic Acid Induction-deficient Mutants of Arabidopsis Express PR-2 and PR-S and Accumulate High Levels of Camalexin after Pathogen Inoculation’, The Plant Cell, 11: 1393–404.
Nei, M. (1982) ‘Evolution of Human Races at the Gene Level’. In: Human Genetics, Part A: The Unfolding Genome. New York: Alan R. Liss. 167–81.
Newman, M. E. J. (2006) ‘Modularity and Community Structure in Networks’, Proceedings of the National Academy of Sciences of the USA, 103: 8577–82.
Nicaise, V. et al. (2007) ‘Coordinated and Selective Recruitment of eIF4E and eIF4G Factors for Potyvirus Infection in Arabidopsis Thaliana’, FEBS Letters, 581: 1041–6.
Pagán, I., Alonso-Blanco, C., and García-Arenal, F. (2008) ‘Host Responses in Life-history Traits and Tolerance to Virus Infection in Arabidopsis Thaliana’, FPLoS Pathogens, 4: e1000124.
Pagán, I. et al. (2010) ‘Arabidopsis thaliana as a model for the study of plant-virus Co-evolution’, Philosophical Transactions of the Royal Society B, 365: 1983–95.
Parrat, S. R., Numminen, E., and Laine, A. L. (2016) ‘Infectious Disease Dynamics in Heterogeneous Landscapes’, Annual Review of Ecology, Evolution, and Systematics, 47: 283–306.
Pfenning, K. (2001) ‘Evolution of Pathogen Virulence: The Role of Variation in Host Phenotype’, Proceedings of the Royal Society B Biological Sciences, 268: 755–60.
Pieterse, C. M. J. et al. (2009) ‘Networking by Small-molecule Hormones in Plant Immunity’, Nature Chemical Biology, 5: 308–16.
——— et al. (2012) ‘Hormonal Modulation of Plant Immunity’, Annual Review of Cellular and Developmental Biology, 28: 489–521.
Regoes, R. R., Nowak, M. A., and Bonhoeffer, S. (2000) ‘Evolution of Virulence in a Heterogeneous Host Population’, Evolution, 54: 64–71.
Rodríguez, D., and Torres-Sorando, L. (2001) ‘Models of Infectious Diseases in Spatially Heterogeneous Environments’, Bulletin of Mathematical Biology, 63: 547–71.
Rozas, J. et al. (2017) ‘DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets’, Molecular Biology and Evolution, 34: 3299–302.
RStudio Team (2020) RStudio: Integrated Development for R. Boston, MA: RStudio, PBC.
Sallinen, S. et al. (2020) ‘Intraspecific Host Variation Plays a Key Role in Virus Community Assembly’, Nature Communications, 11: 5610.
Schmid-Hempel, P. (2013) ‘Host-parasite genetics’. In: Oxford University Press (ed) Evolutionary Parasitology: The Integrated Study of Infections, Immunology, Ecology, and Genetics. Chapter 10. pp. 259–72. Oxford, UK.
Schmid-Hempel, P., and Koella, J. C. (1994) ‘Variability and Its Implications for Host-parasite Interactions’, Parasitology Today, 10: 98–102.