Evolution 2014, **68**(9): 2467-2480.

DOI: 10.1111/evo.12458

EXPERIMENTAL EVOLUTION OF AN EMERGING PLANT VIRUS IN HOST GENOTYPES THAT DIFFER IN THEIR SUSCEPTIBILITY TO INFECTION

Julia Hillung[1], José M. Cuevas[1,2], Sergi Valverde[3,4], and Santiago F. Elena[1,5,6]

[1]*Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV), Campus UPV CPI 8E, C/ Ingeniero Fausto Elio s/n, 46022 València, Spain*

[2]*Current address: Instituto Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, C/ Catedrático Agustín Escardino 9, Paterna, 46980 València, Spain*

[3]*Complex Systems Laboratory (ICREA-UPF), Parc de Recerca Biomèdica Barcelona, C/ Dr. Aiguader 88, 08003 Barcelona, Spain*

[4]*Institut de Biologia Evolutiva (CSIC-UPF), C/ Passeig Marítim de la Barceloneta 37, 08003 Barcelona, Spain*

[5]*The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

[6]*E-mail: sfelena@ibmcp.upv.es*

Running title: TEV adaptation to different *A. thaliana* ecotypes

Keywords: Emerging viruses; Experimental evolution; Generalists; Specialists; Virus evolution

Word count: 7318; Tables: 3; Figures: 4

This study evaluates the extent to which genetic differences among host individuals from the same species conditions the evolution of a plant RNA virus. We performed a three-fold replicated evolution experiment in which *Tobacco etch potyvirus* isolate *At*17b (TEV-*At*17b), adapted to *Arabidopsis thaliana* ecotype L*er*-0, was serially passaged in five genetically heterogeneous ecotypes of *A. thaliana*. After 15 passages we found that evolved viruses improved their fitness, showed higher infectivity and stronger virulence in their local host ecotypes. The genome of evolved lineages was sequenced and putative adaptive mutations identified. Host-driven convergent mutations have been identified. Evidences supported selection for increased translational efficiency. Next, we sought for the specificity of virus adaptation by infecting all five ecotypes with all 15 evolved virus populations. We found that some ecotypes were more permissive to infection than others, and that some evolved virus isolates were more specialist/generalist than others. The bipartite network linking ecotypes with evolved viruses was significantly nested but not modular, suggesting that hard to infect ecotypes were infected by generalist viruses whereas easy to infect ecotypes were infected by all viruses, as predicted by a gene-for-gene model of infection.

Species genetic diversity influences the dynamics of ecosystems (Loreau et al. 2011). Large within-species diversity means more productive ecosystems, more resilience against perturbations and faster recovery after disturbances (Reusch et al. 2005). Parasites are one of the most common perturbing factors of ecosytems (Poulin 1998), and are thus a factor to consider when studying the effects of host genetic diversity. Epidemics can lead to population extinctions (Pounds et al. 2006; Rauch and Weisser 2006), but even without becoming epidemic, parasites may reduce host density, growth and productivity and thus affect the functioning of ecosystems (Pounds et al. 2006). Furthermore, host genetic diversity for genes involved in resistance against infection influence the evolution of parasites as well (Kessing et al. 2006; Altermatt et al. 2008).

Host populations with low genetic diversity in resistance-related loci show higher infection prevalence than populations with great diversity. Evidences first come from agriculture, where genetically homogeneous crops were shown to be more susceptible to diseases than crops grown in heterogeneous mixtures (Mundt 2002). Additional support has been accumulating along years for bacteria (Dennehy et al. 2007), arthropods (Calleri et al. 2006; Altermatt and Ebert 2008; Reber et al. 2008; Ganz and Ebert 2010), mollusks (Grosholz 1994; Webster and Woolhouse 1998), fishes (Ferguson and Drahushchak 1990; Arkusch et al. 2002), birds (Reid et al. 2003; Whiteman et al. 2005), and mammals (Burgner et al. 2006; Tibayrenc 2007; Luikart et al. 2008; Capparelli et al. 2009). At one extreme, populations formed only by susceptible individuals have a high chance of being extinct. At the other extreme, populations composed of only resistant individuals will have the lowest infection rate. However, natural populations are composed by individuals with different degrees of susceptibility. The question of how host genetic diversity influences the evolution of microparasites is the topic of this study.

Since microparasites evolve faster than their multicellular hosts, chances of generating escape mutants while replicating in a permissive host genotype are high, thus jeopardizing the viability of populations formed by individuals resistant only to the non-mutated parasite (Martínez et al. 2012). Indeed, this process may lead to local adaptation of parasites, where they have higher fitness in its current host but lower in alternative ones (Kaltz and Shykoff

1998; Bedhomme et al. 2012). Local adaptation to a particular host genotype reduces the likelihood of successful transmission to a different one, thus slowing down the rate of epidemic spread. While heterogeneity in susceptibility slows down the spread of infection, too much diversity will allow infection by a wider range of parasites (Van Baalen and Beekman 2006). Therefore, the long-term outcome of the interplay between host and virus populations depends on the degree of genetic diversity in both contenders.

The interaction between host genotypes and parasite genotypes have been modeled in the context of two different approaches. At the one extreme, the gene-for-gene (GFG) model, where a parasite genotype can infect all host genotypes and a universally susceptible host genotype exists (Flor 1956). Resistance occurs when a host "resistance" gene is matched by at least one parasite "avirulence" gene. Polymorphism in infectivity and resistance can be maintained only if virulence pays a cost. At the opposite side, the matching alleles (MA) model is based on self *vs.* non-self recognition systems in invertebrates. Infection is not possible unless the parasite 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. The classic approach to study these models has been to test a number of host and parasite genotypes in a cross-infection experiment and use ANOVA techniques to evaluate if a significant host by parasite interaction exists. In recent years networks theory have been applied to bacteria-by-phage infection matrices (Flores et al. 2011, 2013; Weitz et al. 2013). This approach has revealed (*i*) that infection networks show a characteristic nested structure caused by the existence of generalist phages that infect most bacteria and very permissive bacteria available to most phages. (*ii*) Networks are anti-modular, as groups of phages tend to infect non-overlapping groups of hosts. GFG predicts infectivity matrices to be nested, since the host range of specialist virus is a subset of the host range of generalist viruses. By contrast, MA predicts infection matrices to be modular, since infection is likely for viruses and host from the same module but rare for those belonging to other modules.

Our model pathosystem is composed by *Tobacco etch virus* (TEV; genus *Potyvirus*, family Potyviridae) and *Arabidopsis thaliana* (Agudelo-Romero et al. 2008a, 2008b; Lalić et al. 2010;

Hillung et al. 2012). *A. thaliana* ecotypes vary in susceptibility to TEV (Mahajan et al. 1998): some allow long-distance movement from inoculated to non-inoculated leaves while other support replication in inoculated leaves but do not allow for systemic movement (Mahajan et al. 1998). Susceptibility depends on the *Restricted TEV Movement* (*RTM*) multigenic system composed of the *RTM1*, *RTM2* and *RTM3* loci (Mahajan et al. 1998; Whitham et al. 1999, 2000; Chisholm et al. 2000, 2001; Cosson et al. 2010a, 2010b). The presence of dominant alleles in all three loci is necessary for resistance, homozygous recessive mutations at any of the three loci results in systemic infection (Chisholm et al. 2000, 2001).

Agudelo-Romero et al. (2008b) performed an evolution experiment in which TEV was adapted to ecotype L*er*-0 (Table 1). The ancestral TEV systemically infected L*er*-0 plants, although the infection was asymptomatically. After 17 passages, the resulting strain, TEV-*At*17, fixed six point mutations, improved its accumulation ca. 44-fold and induced severe symptoms. Comparative transcriptomics showed differences between evolved and ancestral viruses: TEV-*At*17 down-regulated developmental and metabolic processes, innate immunity, and responses to abiotic stresses and to infection. Lalić et al. (2010) showed that TEV-*At*17 systemically infected ecotypes that were resistant to the ancestral TEV (Table 1). Furthermore, infectivity, accumulation, and severity of symptoms varied among ecotypes. Hillung et al. (2012) compared the effect of TEV-*At*17 infection on the transcriptome of the five ecotypes listed in Table 1, finding that they differ in the way perceived and responded to infection. L*er*-0, St-0 and Di-2 developed strong symptoms and accumulated large amounts of virus. Ei-2 and Wt-1 developed mild symptoms and accumulated fewer viruses. This classification into two groups also explained the differences in transcriptomic responses among ecotypes: Ei-2 and Wt-1 up-regulated genes involved in abiotic stresses and in the construction of new tissues; the other ecotypes up-regulated defense genes.

In this study we sought to explore whether further evolution of TEV-*At*17 on each ecotype would result in specialization or, by contrast, the new evolved viruses would retain the ability to infect all ecotypes as the starting virus. After a period of experimental evolution, we evaluated the infectivity, virulence and relative fitness of each evolved strain across all five *A. thaliana*

ecotypes and assessed whether significant host genotype by virus strain interactions has emerged. We applied network analyses to evaluate the nestedness and modularity of the infection matrix. The molecular basis of the adaptive process are also explored.

Methods

VIRUS AND PLANT ECOTYPES

A. thaliana L*er*-0 frozen infected material from passage 16 of the evolution experiment described in Agudelo-Romero et al. (2008b) was used to prepare a sap. Fifty 21-days-old L*er*-0 were rub-inoculated with 4 µL of sap containing 10% Carborundum. Plants were maintained in a BSL-2 greenhouse at 16:8 h light:dark and 24:20 ºC day:night until sample collection 21 days post-inoculation (dpi). Symptomatic plants were collected, systemic tissues ground into fine powder and stored at −80 ºC.

The consensus sequence for the whole genome of the viral population was obtained as described in Agudelo-Romero et al. (2008b). The resulting genomic sequence was identical to the one previously reported for TEV-*At*17 (Agudelo-Romero et al. 2008b) with an additional nonsynonymous mutation G6816A (M2224I) in the *NIaPro* cistron. We named this isolate TEV-*At*17b (Hillung et al. 2012).

The five *A. thaliana* ecotypes listed in Table 1 were chosen for this study. According to their genetic makeup, Ei-2, L*er*-0 and St-0 shall be sensitive to infection with the ancestral TEV whereas Di-2 and Wt-1 should not; all are sensitive to TEV-*At*17b (Hillung et al. 2012).

EXPERIMENTAL EVOLUTION

Experimental evolution consisted of three-fold replicated serial passages of TEV-*At*17b in five ecotypes of *A. thaliana* (Table 1). Evolution was initiated as described in Hillung et al. (2012). Passages were carried out every 21 dpi. Infection was confirmed by RT-PCR on upper leafs (Lalić et al. 2010). Randomly chosen infected plants from each ecotype were collected, systemically infected tissues were ground into fine powder in liquid N_2 and used to inoculate the

next generation of plants. Experimental evolution was continued for 15 passages. At the end of experimental evolution five plants from each lineage per ecotype were randomly chosen, homogenized and stored at −80 ºC.

VIRUS GENOMIC RNA PURIFICATION, QUANTIFICATION AND SEQUENCING

RNA extraction from 100 mg tissue per plant was performed using InviTrap® Spin Plant RNA Mini Kit (Invitek) following manufacturer's instructions. The concentration of total plant RNA extracts was adjusted to 50 ng/µL for each sample and the quantification of viral load was done with real time RT-qPCR as described in Lalić et al. (2010). Amplifications were done using an ABI StepOnePlus™ Real-Time PCR System (Applied Biosystems) as follows: 5 min at 42 ºC, 10 s at 95 ºC following 40 cycles of 5 s at 95 ºC and 34 s at 60 ºC. Quantifications were performed in triplicate for each sample.

Full genome consensus sequences of evolved viral isolates were obtained as previously described (Agudelo-Romero et al. 2008b). Chromatogram visualization and contigs assembling were done with Lasergene (DNAStar Inc. Madison WI). Molecular evolutionary genetics analyses were done with MEGA 6.05 (Tamura et al. 2013).

MEASURING INFECTIVITY, VIRULENCE AND FITNESS

To compare the infectivity, virulence and fitness of evolved lineages across host ecotypes, all five ecotypes were inoculated with all 15 evolved viral isolates. To this end infectious sap resulting from the experimental evolution was quantified and diluted to the same concentration. One leaf from 21 days old *A. thaliana* plant was inoculated. Control plants were inoculated only with buffer and maintained in the same conditions. Infection was verified by RT-PCR 21 dpi and infectivity, *I*, was determined as the proportion of infected plants among inoculated plants.

The aerial part of infected and control healthy plants were weighted with a precision of 10 mg 21 dpi. Virulence, *V*, was defined as the reduction in weight due to infection and calculated

as $V = 1 - P/\bar{P}$, where P is the weight of an infected plant and $\bar{P}$ is the average weight of non-inoculated plants from the same ecotype.

Between five and nine infected plants from each ecotype-viral lineage combination were collected. Total RNA of infected plants was extracted and the fraction of viral RNA in total RNA was quantified by RT-qPCR. A Malthusian growth rate per day was computed as $m = \frac{1}{t}\log(C_t/C_0)$, where C_t is the number of TEV genomes/100 ng of total RNA quantified t dpi. Relative fitness was calculated as $W_x = \exp(m_x - m_{\text{TEV-}At17\text{b}})$, where m_x and $m_{\text{TEV-}At17\text{b}}$ are the Malthusian growth rates of the viral isolate x and of the ancestral TEV-*At*17b isolate, respectively, evaluated in the same host.

STATISTICAL ANALYSES

W data were analyzed using GLM (Normal distribution and an identity link function). The model has four random factors, the local host (*LH*), the virus evolutionary lineage (*L*), the test host (*TH*), and the experimental biological replicate (*R*; *i.e.*, an individual plant of the test host inoculated with virus from a lineage evolved in a certain local host). *LH* and *TH* were orthogonal factors, *L* was nested within *LH* (represented as *L*(*LH*)) and *R* was nested within the *TH*×*L*(*LH*) interaction. The model equation was:

$$W_{ijklm} = \mu + LH_i + L(LH)_{ij} + TH_k + (LH\times TH)_{ik} + (TH\times L(LH))_{ijk} + R(TH\times L(LH))_{ijkl} + \varepsilon_{ijklm}, \quad (1)$$

where μ is the grand mean value and ε_{ijklm} is the error associated with individual measure *m*.

I data were analyzed using GLM and the following binary logistic regression equation (Binomial responses and logit link function):

$$\log[I_{ijkl}/(1 - I_{ijkl})] = \mu + LH_i + L(LH)_{ij} + TH_k + (LH\times TH)_{ik} + (TH\times L(LH))_{ijk} + \varepsilon_{ijkl}, \quad (2)$$

where ε_{ijkl} is the error associated with individual measure *l*.

V data were analyzed using GLM (Normal distribution and an identity link function) and the equation:

$$V_{ijkl} = \mu + LH_i + L(LH)_{ij} + TH_k + (LH\times TH)_{ik} + (TH\times L(LH))_{ijk} + \varepsilon_{ijkl}. \quad (3)$$

The magnitude of the different effects included in the models was evaluated using the partial eta-squared (η_P^2) statistic that represents the proportion of the total variability attributable to a given factor. Conventionally, values $\eta_P^2 < 0.05$ are considered as small, $0.05 \leq \eta_P^2 < 0.15$ as medium and $\eta_P^2 \geq 0.15$ as large effects.

The relative extent of genetic divergence versus phenotypic parallelism during experimental evolution is given by the ratio $I_W = \sigma_G(W)/|\overline{\Delta W}|$ (Vasi et al. 1994), where $\sigma_G(W)$ is the between-lineages genetic standard deviation for fitness, and $|\overline{\Delta W}|$ the absolute value of the average change in fitness from the common ancestor. I_W provides a measure of the average genetic difference among lineages relative to the average evolutionary change from the ancestral state. Under the null hypothesis of all phenotypic change being associated to a genetic change, $I_W = 1$.

All statistical analyses were done with IBM SPSS Statistics version 21.

NETWORK STATISTICS

An infection network is considered bipartite when it contains two agents that interact, in this case virus isolates and plant ecotypes. This network can be represented as a Boolean matrix of size $m \times n$ with entries assigned to 1 when there is a reported infection of the pair (virus isolate, plant ecotype) or 0, otherwise. Here m is the number of viral isolates and n is the number of distinct ecotypes. The W matrix was first transformed into a Boolean matrix indicating whether a given viral isolate performs significantly better (1) or worse (0) than the ancestral TEV-*At*17b into a given host genotype.

The nestedness of the infection matrix was calculated using the nestedness temperature calculator algorithm implemented in BINMATNEST (Rodríguez-Gironés and Santamaría 2006). The temperature T of an interaction matrix is estimated by resorting the row order of ecotypes and the column order of viruses such that as many of the interactions occur in the upper left portion of the matrix. T quantifies the extent to which a matrix is perfectly nested ($T = 0$) or if the matrix lacks of any order and elements distribute at random ($T = 100$).

Bipartite networks can be decomposed into disjoint components such that no cross-infections are found between components. The modularity of bipartite infection networks was computed using the standard Bipartite Recursively Induced Modules algorithm (Barber 2007), which uses a local search heuristic to maximize bipartite modularity Q. Q represents how often a particular ordering of virus and ecotypes into modules corresponds to interactions that are primarily inside a module ($Q = 1$ or modular), outside of modules ($Q = -1$ or antimodular) or somewhere in between ($-1 < Q < 1$).

The statistical significance of T and Q was assessed using the general null model proposed by Bascompte et al. (2003). In this model, the probability of each cell being occupied is the average of the probabilities of occupancy of its row and column. Biologically, this means that the probability of drawing an interaction is proportional to the level of generalization (degree) of both the virus isolate and the plant genotype.

Results

EXTENT OF ADAPTATION TO EACH LOCAL HOST

First, we sought to determine whether each evolved viral lineage has increased fitness in its corresponding local host ecotype relative to TEV-*At*17b. W values are shown in Table S1 (gray squares). None of the lineages evolved in Ei-2 or L*er*-0 showed significant increases in W (one-sample t-tests, 1-tail $P \geq 0.058$). Only Di-2/3 showed 0.85% significant increase (one-sample t-test, 1-tail $P = 0.002$). All lineages evolved in St-0 (2.74%, 5.65% and 13.62%, respectively; one-sample t-tests, 1-tail $P \leq 0.031$) and in Wt-1 (1.13%, 1.13% and 2.24%, respectively; one-sample t-tests, 1-tail $P \leq 0.005$) showed large and significant increases.

Next, we sought to explore whether the observed changes in W reflect genetic divergence among independent lineages or instances of parallel phenotypic evolution with no genetic basis. Lineages evolved in Di-2, Ei-2 and L*er*-0 show quite diverse results (Table S1), while lineages evolved in St-0 and Wt-1 show similar increasing trends in their relative fitness. Table 2 shows the maximum likelihood estimates of $\sigma_G(W)$, $|\overline{\Delta W}|$ and I_W for each local host. $I_W > 1$ among

lineages evolved in Di-2, Ei-2 and L*er*-0, indicating parallel phenotypic evolution: differences among independent lineages are small relative to the average change in fitness from the ancestral TEV-*At*17b. However, we could not reject the null hypothesis of an equal contribution of parallelism and divergence ($z \leq 1.277$, $P \geq 0.101$) owed to the large uncertainty associated to estimates of I_W (Table 2). By contrast, $I_W < 1$ for St-0 and Wt-1, ($z \geq 2.196$, $P \leq 0.014$), indicating that the contribution of genetic differences among lineages was more important than average changes in relative fitness.

I had a weaker response to evolution on different hosts than *W* as fewer lineages showed a significant variation, being the trend also variable in sign: lineage St-0/1 reduced its *I* in 83.34% (GLM; $\chi^2 = 9.515$, 1 d.f., $P = 0.002$). By contrast, lineages Di-2/1 (118.55%), Di-2/3 (122.61%), Wt-1/2 (55.55%), and Wt-1/3 (113.88%) (GLMs; $\chi^2 \geq 3.552$, 1 d.f., $P \leq 0.030$) showed significant increases in *I*.

Therefore, we conclude from these analyses that the evolutionary response of TEV-*At*17b depends on the host genotype. Lineages passaged in St-0 and Wt-1 increased fitness in their local hosts by genetic changes that affect fitness, response to Di-2 was weaker and response to Ei-2 and L*er*-0 was null and showing strong phenotypic parallelism.

THE SPECIFICITY OF ADAPTATION

We now turn our attention to explore whether TEV-*At*17b adaptation to ecotypes that differ in susceptibility to infection would come with a fitness cost in the ancestral host L*er*-0. To test this, we run paired *t*-test comparing *W* of lineages in their local host with their *W* in L*er*-0 (Table S1). Lineages evolved in Di-2, Ei-2 and Wt-1 paid no cost in L*er*-0 ($t_3 \leq 2.058$, 1-tail $P \geq 0.075$). By contrast, lineages evolved in St-0 show 8.93% significant fitness cost in L*er*-0 ($t_3 = 3.129$, 1-tail $P = 0.028$).

To test whether adaptation to a novel local host genotype is associated with performance on foreign novel host ecotypes, we compared the *W* values of each evolved lineage on their corresponding local hosts with the values estimated on each of the three new foreign hosts. Results widely varied among local hosts (Table S1 and Fig. 1A). On average, Di-2-evolved

isolates perform better than the ancestor in St-0 and Wt-1 ($t_3 \geq 3.656$, 1-tail $P \leq 0.034$) but not in Ei-2. Ei-2-evolved isolates performed better than the ancestor in St-0 ($t_3 = 3.346$, 1-tail $P = 0.039$) but equally elsewhere. St-0- and Wt-1-evolved isolates were fitter than the ancestor in Ei-2 ($t_3 \geq 2.992$, 1-tail $P \leq 0.048$) but equally fit in all other three hosts. Table 3 shows the results of the GLM analysis described by equation (1). *LH* and *TH* have both highly significant and large effects ($\eta_P^2 > 0.15$ in all cases) on *W*. More interestingly, the interaction between local and test hosts (*LH*×*TH*) also had a large and highly significant effect, thus suggesting that the selective constraints imposed by a given local host affect subsequent performance on unselected hosts. Consistently, lineages are also heterogeneous in their response to their local host (*L*(*LH*)) as well as in their performance across hosts (*TH*×*L*(*LH*)).

Sensu stricto, all lineages have evolved generalists, since adaptation to a local host genotype is always associated to a fitness increase in at least one of the alternative foreign hosts. However, not all host ecotypes have selected for viruses that are equally generalist. A significant negative correlation exists between fitness in local and foreign hosts (Spearman's $r_S = -0.900$, 3 d.f., 1-tail $P = 0.019$). The most permissive hosts (St-0) has selected for the most specialist virus while the less permissive hosts (*e.g.*, Ei-2) have selected for the more generalist viruses.

EVOLUTION OF INFECTIVITY AND VIRULENCE

I data told a similar history than *W*. Data were fitted to the logistic regression model given by equation (2). A highly significant effect has been detected for *LH* (although of medium size; $\eta_P^2 < 0.15$), *TH* as well as for their interaction (in these two cases, of large effect) (Table 3), suggesting that the selective constraints imposed by local host affect subsequent infectivity on foreign hosts. Overall, St-0-evolved lineages show the highest *I* values in their local host, while Ei-2-evolved lineages show the lowest *I* in Ei-2 (Fig. 1B). Likewise, evolved lineages have higher levels of *I* in Ei-2 and lower in L*er*-0.

V data were fitted to model equation (3). A highly significant effect has been detected for *LH*, *TH* as well as for their interaction (the size of *LH* and *TH* effects was medium but large for

the interaction) (Table 3), suggesting that the virulence of the evolved isolates depends, in a non-additive manner, both on the local host ecotype wherein a lineage has evolved as well as on the host ecotype in which *V* has been measured. Lineages evolved in Di-2 and St-0 are less virulent, while lineages evolved in Ei-2 are the most virulent (Fig. 1C). St-0, L*er*-0, Wt-1, and Di-2 plants show more aggressive symptoms regardless the viral isolate inoculated, whereas Ei-2 tends to be more tolerant to infection with all evolved strains (Fig. 1C). Fig. 1C also illustrates that the more virulent isolates in their local host are less virulent they are in their alternative hosts.

ANALYSIS OF INFECTION NETWORKS

The modularity of the bipartite network shown in Fig. 2B was low (0.192) and not significantly different from the null expectation ($E(Q)$ = 0.222±0.049, P = 0.202). The infection matrix shown in Fig. 2A has a temperature of T = 13.300, a value that is significantly lower than expected by the null model ($E(T)$ = 29.343±8.487, P = 0.019), meaning the infection matrix is significantly nested.

The first row in the matrix corresponds to lineage L*er*-0/2 (Fig. 2A). This is the only lineage whose fitness across all host ecotypes is not different from that of the ancestral virus, not surprising since it has no mutations that make it different from TEV-*At*17b (see below). All other isolates are worse than the ancestral isolate at least in one host ecotype. At the one extreme, Di-2/3 and Ei-2/2 are worse than the ancestral only in Ei-2; at the other extreme, St-0/1, St-0/3 and Wt-1/2 are worse than the ancestral in all ecotypes but one. The first column in the infection matrix (Fig. 2A) corresponds to ecotype St-0; it was the most susceptible one, as it was infected by 13 evolved viral isolates better than by the ancestral virus. At the other side, the most reliant ecotype was Ei-2, which was only successfully infected by two of the evolved viral isolates.

Therefore, we conclude that during our evolution experiment, the matrix of virus-host interaction evolved no modularity but significant nestedness, as predicted by GFG.

GENOMIC EVOLUTION

All lineages except L*er*-0/2 contain at least one mutation. A total of 79 independent mutations occurring at 62 different nucleotide sites were identified with a range of two to eight mutations per lineage (Fig. 3 and Table S2). Forty-three mutations were synonymous and 36 nonsynonymous. Twenty-four mutations were not unique and of the 62 polymorphic sites identified, eight were mutated in multiple independent lineages. Three of these not unique mutations were exclusive of Ei-2 lineages (C2116U, G3639A and G6420A). Mutation C795U was shared by St-0 lineages. Mutation G1272U was shared by Di-2 and St-0 lineages. Mutation A9240G was common to lineages L*er*-0/1, St-0/1 and Wt-1/2. These six convergent mutations were synonymous. Nonsynonymous mutation C2912A (A923D in *P3* cistron and L923I in the overlapping *P3N-PIPO* cistron) was shared by lineages Di-2/1 and St-0/1. Nonsynonymous mutation C8636U (S2831L in CP) was shared by all Di-2- and L*er*-0-evolved lineages and by Wt-1/3. In addition, lineages Ei-2/2 (C8624U) and Ei-2/3 (U8623C) each has a nucleotide substitutions affecting the same codon at *CP* cistron but resulting in different amino acid replacements (S2827L and S2827P). These convergent mutations are reflected in the clustering pattern shown in a ML tree, where more diverse clusters alternate with clusters defined by a common host (Fig. S1). Synonymous and nonsynonymous mutations distributed evenly among shared and unique mutations (Fisher's exact test: $P = 0.260$).

Treating each lineage as an observation and each host ecotype as a subpopulation, the average nucleotide diversity within host is $\hat{\pi}_S = 0.167 \pm 0.008$ (±1 SD; 1000 bootstrap samples). On the other hand, the nucleotide diversity for the entire sample is $\hat{\pi}_T = 0.187 \pm 0.014$. Therefore, the estimate of inter-host nucleotide diversity is $\hat{\delta}_{ST} = 0.019 \pm 0.010$, and thus the estimate of the proportion of inter-host nucleotide diversity, known as coefficient of nucleotide differentiation (Nei 1982), is $\widehat{N}_{ST} = \hat{\delta}_{ST}/\hat{\pi}_T = 0.103 \pm 0.043$, a value significantly greater than zero ($z = 2.395$, 1-tail $P = 0.004$). Thus, we conclude that minor yet significant genetic differentiation has been generated among viruses replicating in different host genotypes. To assess whether selection played a role in genetic differentiation among host genotypes, we

performed a *D* test (Tajima 1989) and found that it was significantly negative ($D = -2.172$, $P = 0.015$). The significance of this finding will be discussed later.

ASSOCIATION BETWEEN MOLECULAR DIVERSITY AND GENETIC VARIANCE FOR FITNESS

A test for adaptive evolution could be done by looking at the correlation between genomic diversity of evolved viruses within each local host genotype and $\sigma_G(W)$. If genomic diversity is neutral then we expect no correlation between these two traits. By contrast, if genetic differences among lineages translate in differences in *W*, then we expect a significant positive correlation. To perform this test, we first computed the mean genetic diversity among lineages within each local host (Fig. 4A). We found the lowest genetic diversity for the Ei-2 lineages and the largest one for the St-0 lineages. Fig. 4B shows the association between genetic diversity and $\sigma_G(W)$. The expected positive correlation was not observed for the entire dataset (Spearman $r_S = 0.400$, 3 d.f., 1-tail $P = 0.253$). However, if the data point corresponding to Wt-1 is removed from the analysis, the correlation becomes significant ($r_S = 1.000$, 2 d.f., 1-tail $P < 0.001$). This result suggests that, with the exception of lineages evolved in Wt-1, genomic differences among lineages evolved in a given local host explain the amount of genetic differences for relative fitness among lineages. Therefore, we conclude that some of the mutations observed are beneficial and explain differences in fitness among lineages in their local host genotype.

SELECTION FOR TRANSLATIONAL EFFICIENCY AT SYNONYMOUS SITES

A possible explanation for convergence at synonymous sites is that selection for translational efficiency would result in the replacement of poorly used codons by synonymous ones for which the host cell has a large pool of tRNAs. Table S2 includes the ancestral and mutated codons as well as the frequency of usage (*f*) for *A. thaliana*. For each of the 62 mutations, we computed the relative change in usage between the evolved and the ancestral codons $C =$

$f_{evolved}/f_{ancestral} - 1$. Values of $C > 0$ means that the mutation transforms a codon into a more used one whilst $C < 0$ values imply that mutations go to a more rare codon.

If the hypothesis of selection operating at the translation level is true, then we expect the $\bar{C}$ for convergent synonymous mutations to be positive and significantly larger than the value observed for all other types of mutations. For convergent synonymous mutations $\bar{C}$ = 0.494±0.351 (±1 SEM), whereas for the rest of mutations was $\bar{C}$ = −0.009±0.066, being the difference between groups significant (two-samples *t*-test: t_{60} = 2.185, 1-tail P = 0.016). Therefore, we conclude that convergent synonymous mutations fixed during evolution resulted in codons that were ~50% more used by the *A. thaliana* translational machinery than the original ancestral codon.

Discussion

HETEROGENEITY IN HOST SUSCEPTIBILITY AND THE EVOLUTION OF SPECIALIST AND GENERALIST VIRUSES

The evolution of host range in RNA viruses has received considerable attention due to its implication in emerging infectious diseases. Plant viruses have highly variable host ranges: some are specialists infecting only one or few related species while others are generalists that infect a wide range of hosts from different taxonomic groups. Most previous studies on the evolution of plant virus host range explored the effect on fitness traits of passaging viruses into a single host species (Yarwood 1970; Rico et al. 2006; Agudelo-Romero et al. 2008b, 2008c; Wallis et al. 2007; Bedhomme et al. 2012) or alternating between two host species (Bedhomme et al. 2012). Antagonistic pleiotropy, *i.e.*, evolved viral isolates perform worse in the original host than their ancestors, is common. Antagonistic pleiotropy limits the range of adaptation and promotes the evolution of ecological specialization (Remold 2012). Another observation, extensible to animal viruses evolving in cell cultures, is that frequent alternation between host species results in generalist viruses with increased fitness in all the alternative hosts with no apparent fitness cost (Remold et al. 2008; Bedhomme et al. 2012).

Despite this obvious interest in among-species transmission, scant attention has received the effect of genetic differences within species in the evolution of plant RNA viruses. This lack of studies is surprising given the importance of host genetic diversity in the emergence, spread and prevalence of infectious diseases and that the earliest evidence for such effect on prevalence come from the field of agronomy (Mundt 2002). Here, we brought the studies of plant virus host range evolution one step further by analyzing the effect of genetic differences among ecotypes of *A. thaliana* in the evolution of an emerging virus. Genetic differences among ecotypes created new challenges for the virus. St-0 and Wt-1 are the ecotypes for which the locally evolved lineages show a larger increase in fitness. Not surprisingly, the contribution of genetic divergence *vs*. parallelism is stronger in St-0 and Wt-1 than in the other three ecotypes.

An unexpected observation is that the magnitude of the fitness improvements in the local hosts was not dependent on the genetic makeup of the *RTM* loci. Since only ecotypes carrying *rtm* alleles were susceptible to infection with the wildtype TEV strain, one would expect more room for fitness improvement in Di-2 and Wt-1. However, this was not the case. A possible explanation is that upon adaptation to L*er*-0, TEV-*At*17b acquired the capacity to systemically infect L*er*-0, surpassing the *RTM*-mediated resistance and thus differences in these three loci do not represent a constraint anymore.

Hillung et al. (2012) classified *A. thaliana* ecotypes into two groups depending on their transcriptomic response to TEV-*At*17b infection. In any case, incorporating these two groups as a factor in the linear models did not result in improvements in explanatory power (*e.g.*, for *W* data, BIC = −7137.105 for the model here presented and BIC = −4101.527 for the model incorporating an additional factor), so we concluded that the outcome of the evolution experiments was not affected by the differences in symptoms or in transcript profiles of the ecotypes. Therefore, the observed differences in viral fitness among isolates may represent adaptation to one or, more likely, many of the small differences among the genetic or biochemical components of the ecotypes.

The analysis of the infection matrix allows to conclude that adaptation to Di-2 and Ei-2 comes with no fitness cost in the ancestral host ecotype L*er*-0, while passages in St-0 and Wt-1

come with a significant cost, suggesting that the targets of adaptation in Di-2 and Ei-2 are more similar to those for L*er*-0 while they may be different for St-0 and Wt-1. When the fitness of evolved lineages was tested across all five ecotypes, we found that Di-2-evolved lineages showed higher fitness across all hosts than the rest of lineages. In this sense, Di-2-evolved lineages can be considered as the most generalist ones, whereas Ei-2-evolved lineages are the most specialized. It is usually assumed that generalism cannot evolve in the presence of fitness tradeoffs across hosts and thus a specialist on a given host will always be able to outcompete a generalist sharing that host. As a consequence, there would be no single genotype that has the highest fitness in all environments. However, Remold (2012) presented a model to explain the evolution of specialists and no-cost generalists: epistatic pleiotropy. Epistatic pleiotropy occurs when viral genetic backgrounds differ in how the effect of an allele depends on the host. Under epistatic pleiotropy, viral populations may achieve either specialism or no-cost generalism, depending on the host in which they evolve, despite the existence of tradeoffs. In agreement with these expectations, we found that no isolate was superior across all five ecotypes and some isolates paid a fitness cost whereas other did not. Previous data have shown that reciprocal sign epistasis is pervasive in TEV genome (Lalić and Elena 2012a) and that epistasis depends on the host species where it is evaluated (Lalić and Elena 2012b). These means specialist lineages would not be able to evolve towards no-cost generalists without crossing a fitness valley.

No evolved viral isolate was superior to all other isolates on every host ecotype. Likewise, we found that no single host ecotype was superior to all others in resistance to every viral isolate (Fig. 2A). Under such conditions it is possible to imagine that host ecotypes and virus gene frequencies would be critical for the evolution of natural populations: selection would favor those viral genotypes able to infect the more susceptible host ecotypes (St-0), with these ecotypes being subsequently disfavored. Thus, frequency-dependent selection may arise in this system.

We should expect modularity in infection networks if host and pathogens preferentially cross-infect within groups (Weitz et al. 2013), while we should expect nestedness if a hierarchy of resistance among hosts and infection ability among viruses exists (Weitz et al. 2013). We

have shown that the evolved infection network (Fig. 2B) was not modular but significantly nested. The lack of modularity can be interpreted as if ecotypes are not similar in their response to the infection of certain viral isolates, so each ecotype responds in a particular manner. A GFG model of interaction between TEV and *A. thaliana* predicts the emergence of nestedness. A GFG mechanism implies that mutations increasing fitness in the new local host exist that do not pay a fitness cost in L*er*-0, thus the set of hosts that an isolate can infect are subsets of each other. At the other side, a MA mechanism implies that by acquiring the ability to infect a given ecotype, viruses may entirely loss the ability to infect L*er*-0. As we discussed above, however, a cost exists for the lineages that have experienced the largest increase in fitness in their local host (St-0 and Wt-1), but not for the lineages that show minor fitness increases in their local hosts (Di-2 and Ei-2). Therefore, our data do not fully match to any of these two extreme models but to some intermediate mechanism by which TEV isolates have evolved the ability to infect new hosts and lose their infectivity in the original host ecotype only under certain conditions.

TRADEOFFS BETWEEN FITNESS, VIRULENCE AND INFECTIVITY

Wrapping up the results from the three traits measured, lineages evolved in Di-2 reached the highest fitness across all host genotypes, although such fitness increase was not paralleled by increases in infectivity or in virulence. By contrast, Ei-2-evolved lineages had the lowest fitness across all hosts, being the less infectious but the more virulent ones. St-0-evolved lineages were the most infectious ones but at the same time the less virulent. Indeed, a negative significant correlation exists between average virulence and infectivity ($r_S = -0.900$, 3 d.f., $P = 0.037$). From the perspective of permissiveness to infection, on average, the highest infectivity has been observed for Ei-2. This genotype also develops the weakest symptoms.

Provided that virulence does not represent any clear advantage for the parasite, explaining why most parasites induce symptoms in their hosts is a relevant question. A common assumption is that virulence is an unavoidable consequence of parasite's multiplication (Lenski and May 1994) and thus a positive association must exist between virulence and accumulation.

Such association has been previously reported for TEV infecting pepper (Agudelo-Romero et al. 2008c) and for *Cauliflower mosaic virus* (CaMV) infecting turnip (Doumayrou et al. 2012). Here, we failed to find this association. Likewise, in a previous study this association was not found for TEV genotypes that differed in single point mutations and evaluated in their natural host tobacco (Carrasco et al. 2007). This apparent contradiction suggests that the positive association may be pathosystem-dependent. Two reasons can explain lack of positive correlation. First, virulence estimates are too noisy for reliable statistical inferences. In this sense, most of the observed phenotypic variance (71.9%) was not explained by genetic differences among isolates but attributable to noise. Second, a correlation does not exist and many other factors influencing the progression of viral infection would explain virulence. In particular, virulence would not depend on within-host replication if the extent of damage is not proportional to the amount of viral particles, as in the case of a hypersensitive response (Morel and Dangl 1997), if expressing the systemic acquired resistance pathway is costly (Heidel et al. 2004), or if allocating resources to defense detracts from vegetative growth or reproductive effort (Heil 2001; Pagán et al. 2008).

An adaptive explanation for the evolution of virulence is the tradeoff hypothesis that proposes that virulence must positively correlate with transmission (Anderson and May 1982). Depending on the form of the virulence-transmission function, the tradeoff hypothesis predicts that virulence may evolve either to maximal levels or to an intermediate optimum. The latter occurs when the costs of virulence are not outweighed by the benefits of additional increase in transmission (Frank 1996). In the case of plant virus, a positive correlation between transmission and virulence was found for CaMV (Doumayrou et al. 2012). In sharp contrast, we found a significant negative correlation between virulence and infectivity. During our artificial transmission experiment, lineages were transmitted mechanically regardless their virulence, thus we may have relaxed the tradeoff resulting in independent evolution of virulence and infectivity.

ON THE MOLECULAR BASIS OF ADAPTATION

The characterization of the full genome consensus sequences of the evolved isolates revealed some interesting features. First, evolved TEV isolates contained a variable number of nucleotide substitutions, including both synonymous and nonsynonymous. Second, significant genetic differentiation among lineages was generated during our evolution experiment. We evaluated whether this divergence was due to the action of natural selection and found a significant and negative Tajima's *D* value. $D < 0$ values are compatible with three explanations: (*i*) segregation of slightly deleterious mutations, (*ii*) purifying selection and (*iii*) a fast population expansion. In an expanding population, new mutations may be segregating and will be observed as singletons. In our case, we have observed 55 singletons that have inflate the number of segregating sites and cause $D < 0$.

We have other evidences supporting adaptive evolution at the genomic level. First, we observed significant and host ecotype specific changes in relative fitness. Second, we observed a significant positive correlation between genomic diversity and $\sigma_G(W)$. This positive association between genomic diversity and differences in fitness among lineages is explained by the adaptive value of some of the mutations. Third, supporting this, we observed a number of convergent mutations, many of which appear to be dependent on the host ecotype. And fourth, we observed a significant genomic differentiation of viral populations evolving on different host genotypes.

While convergent evolution at nonsynonymous sites is explained as a consequence of identical selective pressures and the existence of limited accessible adaptive pathways, molecular convergence at synonymous sites is more problematic to explain. Convergent evolution at synonymous sites has been often observed in experimental evolution of RNA viruses (Bull et al. 1997; Wichman et al. 1999; Cuevas et al. 2002; Novella et al. 2004; Remold et al. 2008; Acevedo et al. 2013; Cabanillas et al. 2013) as well as in evolution of resistance to antiviral drugs (Nijhuis et al. 1999; Martínez-Picado et al. 2000), although such observations are still scarce for plant viruses (*e.g.*, Lafforgue et al. 2011). Furthermore, studies analyzing the mutational landscapes of RNA viruses have shown significant fitness effects associated to silent mutations (Sanjuán et al. 2004; Carrasco et al. 2007; Cuevas et al. 2012; Acevedo et al. 2013)

thus supporting the notion that, at least for RNA genomes, equating synonymous substitution with neutral substitution is not always valid. Three non-mutually exclusive explanations can be brought forward to explain convergent evolution at synonymous sites: (*i*) the necessity to preserve regulatory secondary RNA structures, (*ii*) preventing the formation of long dsRNA structures that may be targets of RNA silencing, and (*iii*) selection for translational efficiency would result in the replacement of poorly used codons by synonymous ones for which the host cell has a large pool of tRNAs. Without discarding hypotheses (*i*) and (*ii*), our results suggest that selection for translational efficiency may, in part, explain convergent synonymous mutations.

Conclusions

Most evolution experiments with plant viruses seeking for the evolution of generalist and specialist viruses employed different host species. In this study we explored the effect that within-species genetic variability for susceptibility has in virus adaptation. We found that some plant ecotypes selected for more generalist viruses whereas other ecotypes selected for more specialist viruses. We found that permissive hosts selected for specialist viruses while restrictive hosts selected for generalist viruses. No evolved virus was superior to all others in every ecotype and no ecotype was resistant against all evolved virus. Such nestedness of the infection matrix creates the conditions for frequency-dependent selection to operate in the long-term.

Despite the similarity among hosts and the relatively short evolution time, independent lineages have accumulated genetic variation and have diverged from each other. The genomic characterization of the evolved lineages has shown cases of host-dependent convergent mutations, including some synonymous cases that may contribute to improve translational efficiency.

Finally, the relative fitness of evolved strains is independent from virulence and infectivity. By contrast, virulence and infectivity are linked, as more virulent viruses are transmitted worse

than temperate ones. This tradeoff and the lack of correlation between virulence and relative fitness have important implications for the evolution of virulence in this pathosystem.

ACKNOWLEDGMENTS

The authors thank Francisca de la Iglesia and Paula Agudo for excellent technical assistance. We acknowledge grant BFU2012-30805 from the Spanish Ministerio de Economía y Competitividad to SFE. JMC was supported by a JAE-doc contract from CSIC. JH was supported by a pre-doctoral fellowship from Ministerio de Economía y Competitividad.

LITERATURE CITED

Acevedo, A., L. Brodsky, and R. Andino. 2013. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature 505:686-690.

Agudelo-Romero, P., P. Carbonell, F. de la Iglesia, J. Carrera, G. Rodrigo, A. Jaramillo, M. A. Pérez-Amador, and S. F. Elena. 2008a. Changes in gene expression profile of *Arabidopsis thaliana* after infection with *Tobacco etch virus*. Virol. J. 5:92.

Agudelo-Romero, P., P. Carbonell, M. A. Pérez-Amador, and S. F. Elena. 2008b. Virus adaptation by manipulation of host's gene expression. PLoS ONE 3:e2397.

Agudelo-Romero, P., F. de la Iglesia, and S. F. Elena. 2008c. The pleiotropic cost of host-specialization in *Tobacco eth potyvirus*. Infect. Genet. Evol. 8:806-814.

Altermatt, F., and D. Ebert. 2008. Genetic diversity of *Daphnia magna* populations enhances resistance to parasites. Ecol. Lett. 11:918-928.

Anderson, R. M., and R. M. May. 1982. Coevolution of hosts and parasite. Parasitology 85:411-426.

Arkush, K. D., A. R. Giese, H. L. Mendonca, A. M. McBride, G. D. Marty, and P. W. Hedrick. 2002. Resistance to three pathogens in the endangered winter-run Chinook salmon (*Oncorhynchus tshawytscha*): effects of inbreeding and major histoincompatibility complex genotypes. Can. J. Fish. Aquat. Sci. 59:966-975.

Barber, M. J. 2007. Modularity and community detection in bipartite networks. Phys. Rev. E 76:066102.

Bascompte, J., P. Jordano, C. J. Melián, and J. M. Olesen. 2003. The nested assembly of plant-animal mutualistic networks. Proc. Natl. Acad. Sci. USA 100:9383-9387.

Bedhomme, S., G. Lafforgue, and S. F. Elena. 2012. Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. Mol. Biol. Evol. 29:1481-1492.

Bull, J. J., M. R. Badgett, H. A. Wichman, J. P. Huelsenbeck, D. M. Hillis, A. Gulati, C. Ho, and I. J. Molineux. 1997. Exceptional convergent evolution in a virus. Genetics 147:1497-1507.

Burgner, D., S. E. Jamieson, and J. M. Blackwell. 2006. Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? Lancet Infect. Dis. 6:653-663.

Cabanillas, L., M. Arribas, and E. Lázaro. 2013. Evolution at increased error rate leads to the coexistence of multiple adaptive pathways in an RNA virus. BMC Evol. Biol. 13:11.

Calleri, D. V., E. M. Reid, R. B. Rosengaus, E. L. Vargo, and J. F. A. Traniello. 2006. Inbreeding and disease resistance in a social insect: effects of heterozygosity on immune-competence in the termite *Zootermopsis angusticollis*. Proc. R. Soc. B 273: 2633-2640.

Caparelli, R., M. Parlato, M. Iannaccone, S. Roperto, R. Marabelli, F. Roberto, and D. Iannelli. 2009. Heterogeneous shedding of *Brucella abortus* in milk and its effect on the control of animal brucellosis. J. Appl. Microbiol. 106:2041-2047.

Carrasco, P., F. de la Iglesia, and S. F. Elena. 2007. Distribution of fitness and virulence effects caused by single-nucleotide substitutions in *Tobacco etch virus*. J. Virol. 81:12979-12984.

Chisholm, S. T., S. K. Mahajan, S. A. Whitham, M. L. Yamamoto, and J. C. Carrington. 2000. Cloning of the *Arabidopsis RTM1* gene, which controls restriction of long-distance movement in *Tobacco etch virus*. Proc. Natl. Acad. Sci. USA 97:489-494.

Chisholm, S. T., M. A. Parra, R. J. Anderberg, and J. C. Carrington. 2001. *Arabidopsis RTM1* and *RTM2* genes function in phloem to restrict long-distance movement of *Tobacco etch virus*. Plant Physiol. 127:1667-1675.

Cosson, P., L. Sofer, Q. H. Le, V. Légar, F. Schurdi-Levraud, S. A. Whitham, M. L. Yamamoto, S. Gopalan, O. Le Gall, T. Candresse, J. C. Carrington, and F. Revers. 2010a. *RTM3* which controls long-distance movement of potyviruses, is a member of a new plant gene family encoding a meprin and TRAF homology domain-containing protein. Plant Physiol. 154:222-232.

Cosson, P., L. Sofer, V. Schurdi-Levraud, and F. Revers. 2010b. A member of a new plant gene family encoding a meprin and TRAF homology (MATH) domain-containing protein is involved in restriction of long distance movement of plant viruses. Plant Signal. Behav. 5:1321-1323.

Cuevas, J. M., P. Domingo-Calap, and R. Sanjuán. 2012. The fitness effects of synonymous mutations in DNA and RNA viruses. Mol. Biol. Evol. 29:17-20.

Cuevas, J. M., S. F. Elena, and A. Moya. 2002. Molecular basis of adaptive convergence in experimental populations of RNA viruses. Genetics 162:533-542.

Dennehy, J. J., N. A. Friedenberg, Y. W. Yang, and P. E. Turner. 2007. Virus population extinction via ecological traps. Ecol. Lett. 10:230-240.

Doumayrou, J., A. Avellan, R. Froissart, and Y. Michalakis. 2012. An experimental test of the transmission-virulence trade-off hypothesis in a plant virus. Evolution 67:477-486.

Ferguson, H. M., and L. R. Drahunshchak. 1990. Disease resistance and enzyme heterozygosity in rainbow trout. Heredity 64:413-417.

Flor, H. H. 1956. The complementary genetic systems in flax and flax rush. Adv. Genet. 8: 29-54.

Flores, C. O., J.R. Meyer, S. Valverde, and J. S. Weitz. 2011. Statistical structure of host-phage interactions. Proc. Natl. Acad. Sci. USA 108:E228-E297.

Flores, C. O., S. Valverde, and J. S. Weitz. 2013. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. ISME J. 7:520-532.

1 Frank, S. A. 1993. Specificity versus detectable polymorphism in host-parasite genetics. Proc.
2 R. Soc. B 254: 191-197.
3 Frank, S. A. 1996. Models of parasite virulence. Q. Rev. Biol. 71:37-78.
4 Ganz, H. H., and D. Ebert. 2010. Benefits of host genetic diversity for resistance to infection
5 depend on parasite diversity. Ecology 91:1263-1268.
6 Grosholz, E. D. 1994. The effect of host genotype and spatial distribution on trematode
7 parasitism in a bivalve population. Evolution 48:1514-1524.
8 Heidel, A. J., J. D. Clarke, J. Antonovics, and X. Dong. 2004. Fitness costs of mutations
9 affecting the systemic acquired resistance pathway in *Arabidopsis thaliana*. Genetics
10 168:2197-2206.
11 Heil, M. 2001. The ecological concept of costs of induced systemic resistance (ISR). Eur. J.
12 Plant Pathol. 107:137-146.
13 Hillung, J., J. M. Cuevas, and S. F. Elena. 2012. Transcript profiling of different *Arabidopsis*
14 *thaliana* ecotypes in response to *Tobacco etch potyvirus* infection. Front. Microbiol.
15 3:229.
16 Kaltz, O., and J. A. Shykoff. 1998. Local adaptation in host-parasite systems. Heredity 81:361-
17 370.
18 Keesing, F., R. D. Holt, and R. S. Ostfeld. 2006. Effects of species diversity on disease risk.
19 Ecol. Lett. 9:485-498.
20 Lafforgue, G., F. Martínez, J. Sardanyés, F. de la Iglesia, Q. W. Niu, S. S. Lin, R. V. Solé, N. H.
21 Chua, J. A. Daròs, and S. F. Elena. 2011. Tempo and mode of plant RNA virus escape
22 from RNA interference-mediated resistance. J. Virol. 85:9686-9695.
23 Lalić, J., P. Agudelo-Romero, P. Carrasco, and S. F. Elena. 2010. Adaptation of *Tobacco etch*
24 *potyvirus* to a susceptible ecotype of *Arabidopsis thaliana* capacitates it for systemic
25 infection of resistant ecotypes. Phil. Trans. R. Soc. B 65:1997-2008.
26 Lalić, J., and S. F. Elena. 2012a. Magnitude and sign epistasis among deleterious mutation in a
27 positive-sense plant RNA virus. Heredity 109:71-77.

Lalić, J., and S. F. Elena. 2012b. Epistasis between mutations is host-dependent for an RNA virus. Biol. Lett. 9:20120396.

Lenski R. E., and R. M. May. 1994. The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. J. Theor. Biol. 169:253-265.

Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. Science 294:804-808.

Luikart, G., K. Pilgrim, J. Visty, V. O. Ezenwa, and M. K. Schwartz. 2008. Candidate gene microsatellite variation is associated with parasitism in wild bighorn sheep. Biol. Lett. 4:228-231.

Mahajan, S. K., S. T. Chisholm, S. A. Whitham, and J. C. Carrington. 1998. Identification and characterization of a locus (*RTM1*) that restricts long-distance movement of *Tobacco etch virus* in *Arabidopsis thaliana*. Plant J. 14:177-186.

Martínez, F., G. Lafforgue, M. J. Morelli, F. González-Candelas, N. H. Chua, J. A. Daròs, and S. F. Elena. 2012. Ultradeep sequencing analysis of population dynamics of virus escape mutants in RNAi-mediated resistant plants. Mol. Biol. Evol. 29:3297-3307.

Martínez-Picado, J., M. P. DePasquale, N. Kartsonis, G. J. Hanna, J. Wong, D. Finzi, E. Rosenberg, H. F. Gunthard, L. Sutton, A. Savara, C. J. Petropoulos, N. Hellmann, B. D. Walker, D. D. Richman, R. Siciliano, and R. T. D'Aquila. 2000. Antiretroviral resistance during successful therapy of HIV type 1 infection. Proc. Natl. Acad. Sci. USA 97:10948-10953.

Morel, J. B., and J. L. Dangl. 1997. The hypersensitive response and the induction of cell death in plants. Cell Death Differ. 4:671-683.

Mundt, C.C. 2002. Use of multiline cultivars and cultivar mixtures for disease management. Annu. Rev. Phytopathol. 40:381-410.

Nei, M. 1982. Evolution of human races at the gene level. Pp. 167-181 *in* B. Bonné-Tamir, ed. Human genetics, part A: The unfolding genome. Alan R. Liss, New York.

Nijhuis, M., R. Schuurman, D. de Jong, J. Erickson, E. Gustchina, J. Albert, P. Schnipper, S. Gulnik, and C. A. B. Boucher. 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS 13:2349-2359.

Novella, I. S., S. Zárate, D. Metzgar, and B. E. Ebendick-Corpus. 2004. Positive selection of synonymous mutations in *Vesicular stomatitis virus*. J. Mol. Biol. 342:1415-1421.

Pagán, I., C. Alonso-Blanco, and F. García-Arenal. 2008. Host responses in life-history traits and tolerance to virus infection in *Arabidopsis thaliana*. PLoS Pathog. 4:e1000124.

Poulin, R. 1998. Evolutionary Ecology of Parasites: From Individuals to Communities. Chapman & Hall, London, U.K.

Pounds, J. A., M. R. Bustamante, L. A. Coloma, J. A. Consuegra, M. P. Fogden, P. N. Foster, E. La Marca, K. L. Masters, A. Merino-Viteri, R. Puschendorf, S. R. Ron, G. A. Sánchez-Azofeifa, C. J. Still, and B. E. Young. 2006. Widspread amphibian extinctions from epidemic disease driven by global warming. Nature 439:161-167.

Rauch, G., and W. W. Weisser 2006. Local and spatial dynamics of a host-parasitoid system in a field experiment. Basic Appl. Ecol. 8:89-95.

Reber, A., G. Castella, P. Christe, and M. Chapuisat. 2008. Experimentally increased group diversity improves disease resistance in an ant species. Ecol. Lett. 11:682-689.

Reid, J. M., P. Arcese, and L. Keller. 2003. Inbreeding depresses immune response in song sparrows (*Melospiza melodia*): direct and indirect-generational effects. Proc. R. Soc. B 270:2151-2157.

Remold, S. K. 2012. Understanding specialism when the jack of all trades can be the master of all. Proc. R. Soc. B 279:4861-4869.

Remold, S. K., A. Rambaut, and P. E. Turner. 2008. Evolutionary genomics of host adaptation in *Vesicular stomatitis virus*. Mol. Biol. Evol. 25:1138-1147.

Reusch, T. B. H., A. Ehlers, A. Hammerli, and B. Worm. 2005. Ecosystems recovery after climatic extremes enhanced by genotypic diversity. Proc. Natl. Acad. Sci. USA 102:2826-2831

Rico, P., P. Ivars, S. F. Elena, and C. Hernández. 2006. Insights into the selective pressures restricting Pelargonium flower break virus genome variability: evidence for host adaptation. J. Virol. 80:8124-8132.

Rodríguez-Gironés, M. A., and L. Santamaría. 2006. A new algorithm to calculate the nestedness temperature of presence-absence matrices. J. Biogeography 33:924-935.

Sanjuán, R., A. Moya, and S. F. Elena. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. USA 101:8396-8401.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30:2725-2729.

Tibayrenc, M. 2007. Human genetic diversity and the epidemiology of parasitic and other transmissible diseases. Adv. Parasitol. 64:377-422.

Van Baalen, M., and M. Beekman. 2006. The cost and benefits of genetic heterogeneity in resistance against parasites in social insects. Am. Nat. 167:568-577.

Vasi, F., M. Travisano, and R. E. Lenski. 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. Am. Nat. 144:432-456.

Wallis, C. M., A. L. Stone, D. J. Sherman, V. D. Damsteegt, F. E. Gildow, and W. L. Schneider. 2007. Adaptation of *Plum pox virus* to a herbaceous host (*Pisum sativum*) following serial passages. J. Gen. Virol. 88:2839-2845.

Webster, J. P., and M. E. J. Woolhouse. 1998. Selection and strain specificity between snails intermediate hosts and their parasite schistosomes. Evolution 52:1627-1634.

Weitz, J. S., T. Poisot, J. R. Meyer, C. O. Flores, S. Valverde, M. B. Sullivan, and M. E. Hochberg. 2013. Phage-bacteria infection networks. Trends Microbiol. 21:82-91.

Whiteman, N. K., K. D. Matson, J. L. Bollmer, and P. G. Parker. 2005. Disease ecology in the Galapagos hawk (*Buteo galapagoensis*): host genetic diversity, parasite load and natural antibodies. Proc. R. Soc. B 273:797-804.

Whitham, S. A., R. J. Anderberg, S. T. Chisholm, and J. C. Carrington. 2000. *Arabidopsis RTM2* gene is necessary for specific restriction of *Tobacco etch virus* and encodes an unusual small heat shock-like protein. Plant Cell 12:569-582.

Whitham, S. A., M. L. Yamamoto, and J. C. Carrington. 1999. Selectable viruses and altered susceptibility mutants in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 96:772-777.

Wichman, H. A., M. R. Badgett, L. A. Scott, C. M. Boulianne, and J. J. Bull. 1999. Different trajectories of parallel evolution during viral adaptation. Science 285:422-424.

Yarwood, C. E. 1970. Reversible host adaptation in *Cucumber mosaic virus*. Phytopathology 60:1117-1119.

Table 1. List of *A. thaliana* ecotypes used in this study and the corresponding resistance phenotype according to the allelic combination at each *RTM* loci relative to the ancestral TEV.

Ecotype	**Origin**	**Genotype**	**Phenotype**
Di-2	France	*RTM1*/*RTM1 RTM2*/*RTM2 RTM3*/*RTM3*	resistant
Ei-2	Germany	*rtm1*/*rtm1 RTM2*/*RTM2 RTM3*/*RTM3*	sensitive
L*er*-0	Germany	*rtm1*/*rtm1 RTM2*/*RTM2 RTM3*/*RTM3*	sensitive
St-0	Sweden	*RTM1*/*RTM1 RTM2*/*RTM2 rtm3*/*rtm3*	sensitive
Wt-1	Germany	*RTM1*/*RTM1 RTM2*/*RTM2 RTM3*/*RTM3*	resistant

1

Table 2. Divergence versus parallelism in relative fitness for lineages evolved independently in the five different *A. thaliana* host genotypes.

Local host	$\lvert\overline{\Delta W}\rvert$	Var($\lvert\overline{\Delta W}\rvert$)	$\sigma_G(W)$	Var($\sigma_G(W)$)	I_W (±1 SEM)
Di-2	2.533×10^{-3}	9.402×10^{-5}	1.198×10^{-4}	2.913×10^{-5}	4.321±56.183
Ei-2	2.300×10^{-3}	1.502×10^{-4}	4.657×10^{-5}	7.016×10^{-9}	2.967±7.121
L*er*-0	3.067×10^{-3}	1.675×10^{-3}	3.105×10^{-4}	9.338×10^{-7}	5.746±80.002
St-0	6.957×10^{-2}	3.173×10^{-3}	2.054×10^{-3}	2.951×10^{-6}	0.652±0.076
Wt-1	1.390×10^{-2}	4.107×10^{-5}	1.106×10^{-5}	4.923×10^{-10}	0.239±0.058

Table 3. GLM analyses of variance for the three traits measured for all evolved lineages across the four new hosts (Di-2, Ei-2, St-0, and Wt-1).

Source of variation	Relative fitness				Infectivity				Virulence			
	Wald's χ^2	d.f.	*P*	η_P^2	Wald's χ^2	d.f.	*P*	η_P^2	Wald's χ^2	d.f.	*P*	η_P^2
Intersection μ	36674175.722	1	< 0.001	1.000	0.000	1	1.000	0.954	6743.107	1	< 0.001	1.000
Local host *LH*	555.603	4	< 0.001	0.318	17.562	4	0.002	0.079	19.098	4	0.001	0.141
Lineage *L*(*LH*)	2667.073	10	< 0.001	0.179	81.834	10	< 0.001	0.667	89.816	10	< 0.001	0.376
Test host *TH*	14629.123	4	< 0.001	0.758	61.624	4	< 0.001	0.550	12.848	4	0.012	0.137
Local host by Test host *LH*×*TH*	4655.737	16	< 0.001	0.278	39.447	16	0.001	0.381	82.540	16	< 0.001	0.357
Test host by Lineage *TH*×*L*(*LH*)	12124.267	40	< 0.001	0.186	86.925	39	< 0.001	1.000	149.934	40	< 0.001	0.161
Biological replicate *R*(*TH*×*L*(*LH*))	53712.912	295	< 0.001	0.979								

Figure 1. (A) Average relative fitness in local and unselected hosts. (B) Average in local and unselected hosts. (C) Average virulence in local and unselected hosts. represents the average value of the three evolved lineages. Error bars correspond t Error intervals for the infectivity of Ei-2 across unselected hosts and for St-0 both i across unselected test hosts expand the entire [0, 1] interval.

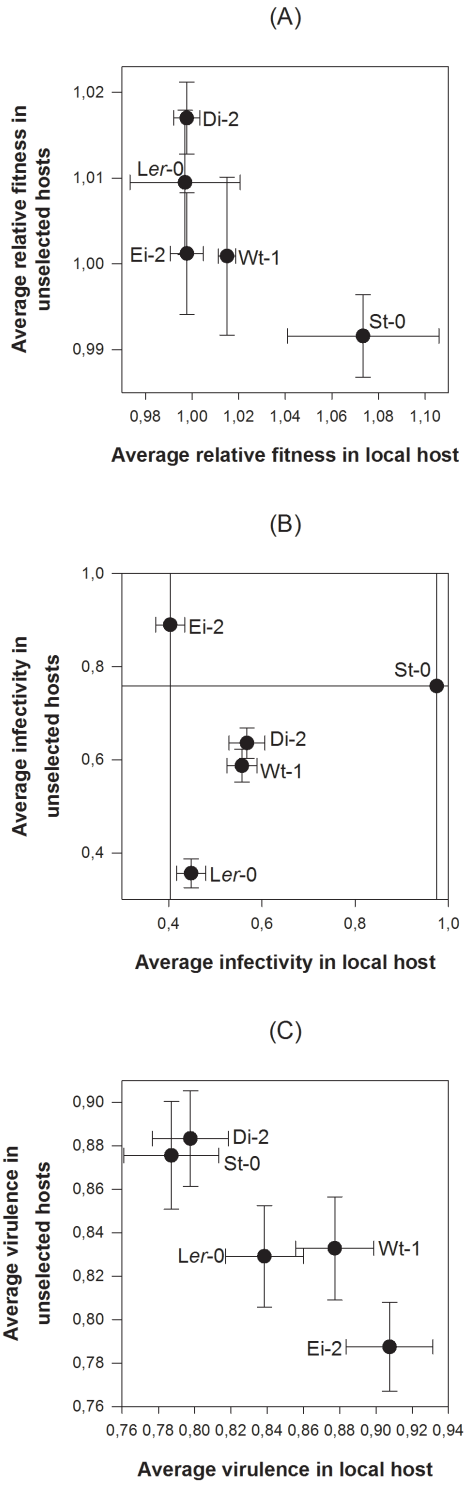


Figure 2. Matrix representation of the host genotype-evolved TEV isolates interactions. (A) The rows represent the 15 viral isolates, and the columns represent the five host genotypes. White cells represent combinations in which the corresponding virus has a relative fitness significantly higher than the ancestral TEV-*At*17b. (B) Bipartite network representation of the infection matrix.

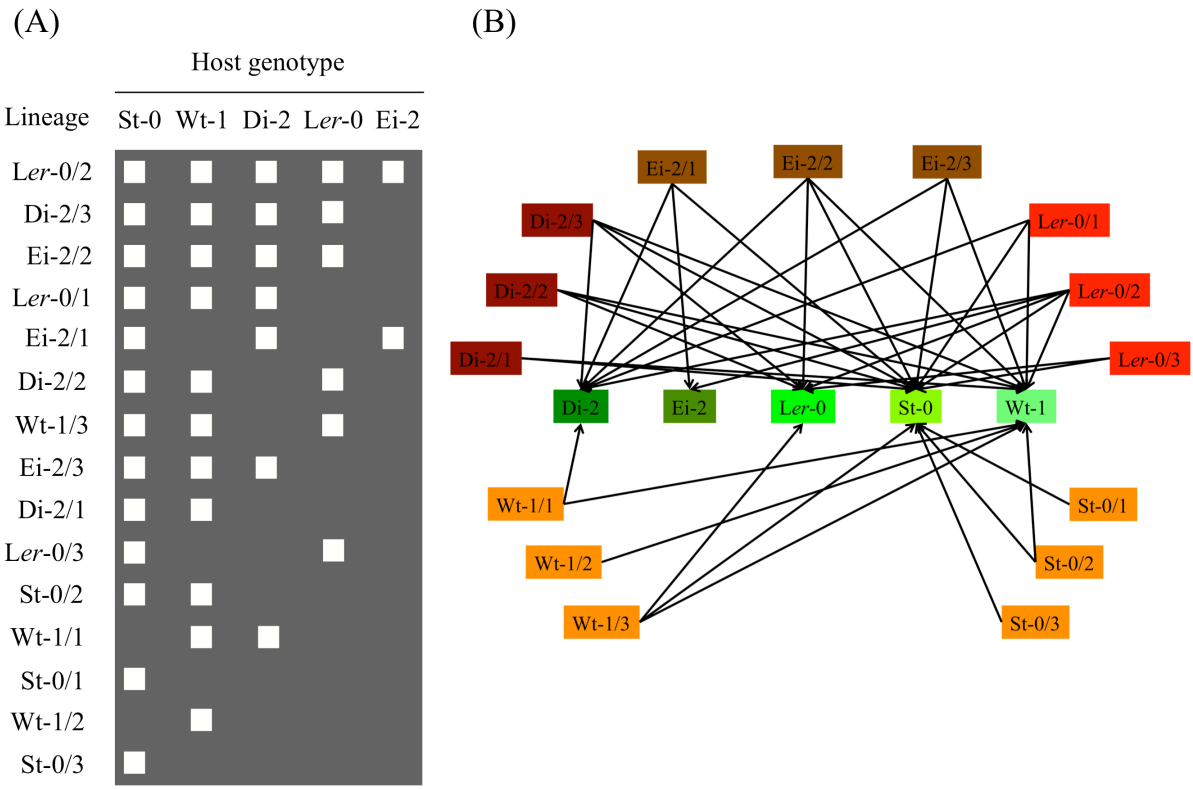


Figure 3. Schematic representation of the collection of mutations observed in 14 of the experimentally evolved lineages (lineage L*er*-0/2 has no mutations). The first line represents the full TEV genome with the position of the 11 mature viral proteins within the ORF. The 14 other lines correspond each to one lineage. Synonymous mutations are represented as empty circles, and nonsynonymous mutations are represented by filled circles. Convergent mutations are indicated above the lineages.

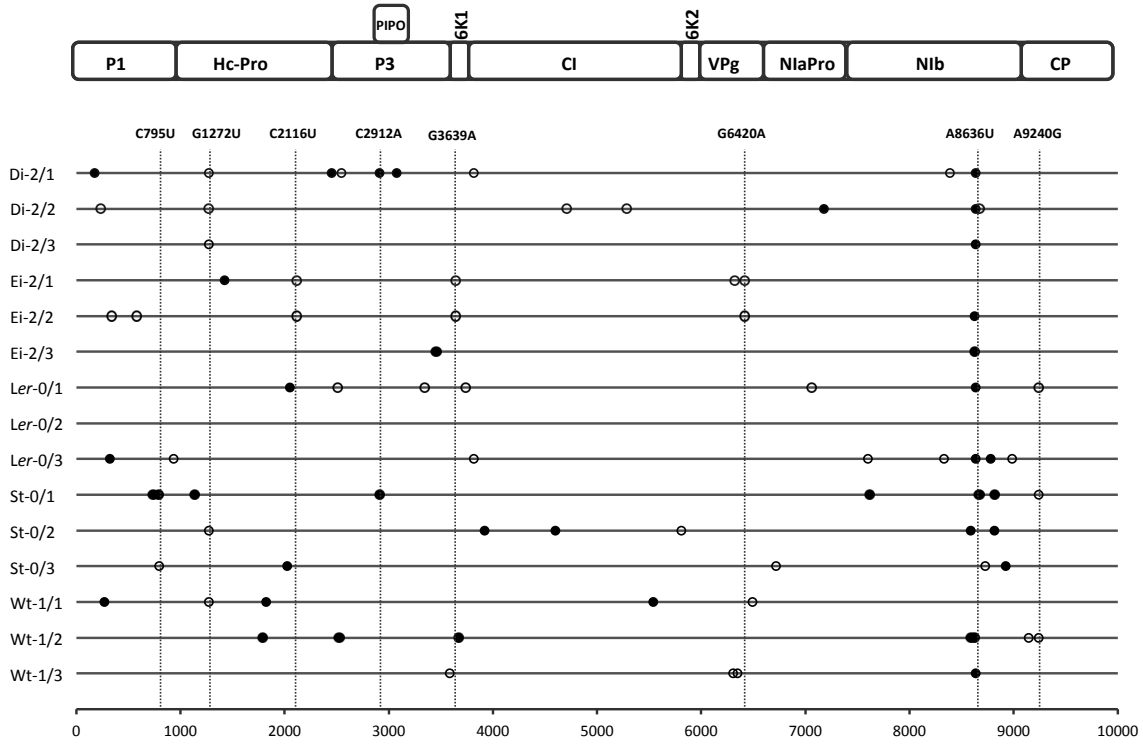
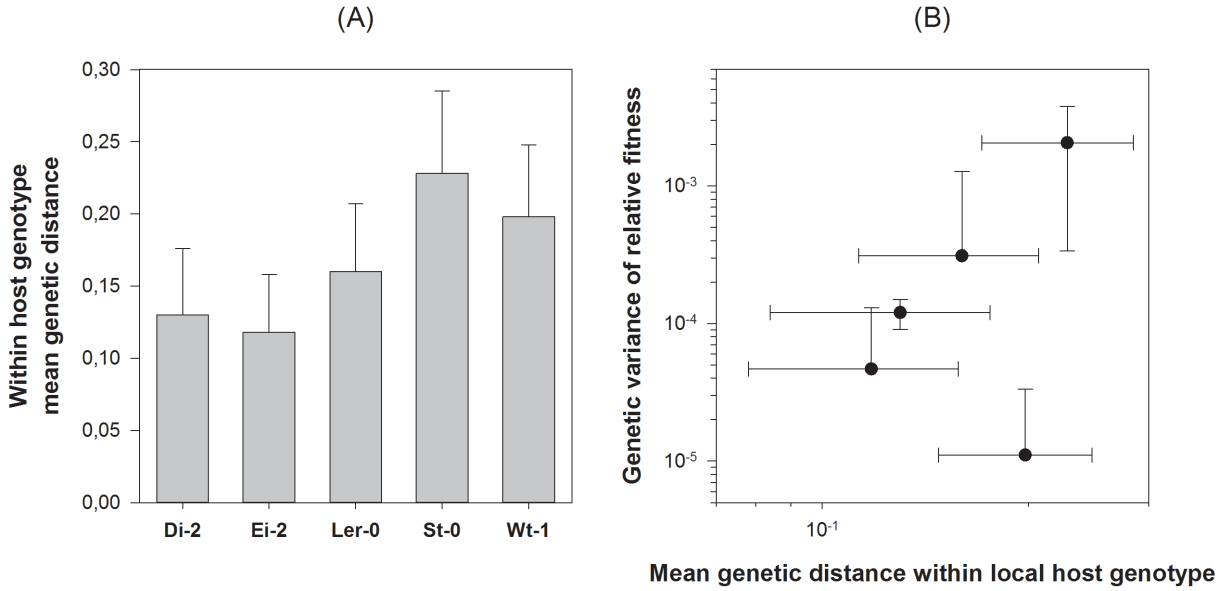


Figure 4. (A) Virus mean genetic diversity within each host genotype. (B) Association between genetic diversity and the genetic component of variance for relative fitness (data from Table 2). Axes in panel B are both in logarithmic scale. All error bars represent ±1 SD.



Supporting Information

The following supporting information is available for this article:

Table S1. Relative fitness values for each evolved lineage measured on each alternative host genotype. The gray shadow indicates tests of adaptation to the local host. Values are the average of a number of infection assays (between five and nine) and errors correspond to ±1 SEM. Asterisks indicate cases in which the value is significantly different from the value estimated for the ancestral TEV-*At*17b isolate (one-sample *t*-tests, $P < 0.05$; significance levels corrected by the FDR method).

Table S2. Complete list of mutations in the 15 independently evolved lineages.

Figure S1. Maximum likelihood phylogenetic tree obtained for the genomic sequences of the evolved lineages. A Kimura 2-parameters model with transitions to transversions rates ratio of 4.99 was the best fitting scheme of nucleotide substitutions; this model was used for constructing this phylogenetic tree as well as for all other molecular analyses reported in the main text. The ancestral sequence TEV-*At*17b was included for rooting purposes. Numbers on the nodes correspond to bootstrap support values. Clusters of isolates evolved in a common host are highlighted with different colors.