

1 **Title:** Multihost experimental evolution of a plant RNA virus reveals local adaptation and
2 host specific mutations.

3

4 **Submission as a research article.**

5

6 **Authors:** Stéphanie Bedhomme¹, Guillaume Lafforgue¹ and Santiago F. Elena^{1,2}

7

8 **Research performed at** *Instituto de Biología Molecular y Celular de Plantas, Consejo*
9 *Superior de Investigaciones Científicas-UPV, Campus UPV CPI 8E, Ingeniero Fausto Elio*
10 *s/n, 46022 València, Spain*

11

12 **Affiliations:** ¹ *Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de*
13 *Investigaciones Científicas-UPV,* ² *The Santa Fe Institute, Santa Fe NM 87501, USA*

14

15 **Author for correspondence:** Stéphanie Bedhomme

16 *Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones*
17 *Científicas-UPV, Campus UPV CPI 8E, Ingeniero Fausto Elio s/n, 46022 València, Spain*

18 Tel: +34 963 877 007 ext 78638 (lab) or 78642 (office)

19 Fax: +34 963 877 859

20 stebed@upvnet.upv.es

21

22 **Keywords:** experimental evolution, local adaptation, *Tobacco etch virus*, parallel evolution,
23 full genome sequence, virulence evolution.

24

25 **Running head title:** multihost viral experimental evolution.

26

27 **Abstract:**

28 For multihost pathogens, simultaneous adaptation to various hosts has important
29 implications for both applied and basic research. At the applied level, it is one of the main
30 factors determining the probability and the severity of emerging disease outbreaks. At the
31 basic level, it is thought to be a key mechanism for the maintenance of genetic diversity both
32 in host and pathogen species. Using *Tobacco etch potyvirus* (TEV) and four natural hosts, we
33 have designed an evolution experiment whose strength and novelty are the use of complex
34 multicellular host organism as hosts, and a high level of replication of the different
35 evolutionary histories and lineages. A pattern of local adaptation, characterized by a higher
36 infectivity and virulence on host(s) encountered during the experimental evolution was found.
37 Local adaptation only had a cost in terms of performance on other hosts in some cases. We
38 could not verify the existence of a cost for generalists, as expected to arise from antagonistic
39 pleiotropy and other genetic mechanisms generating a fitness trade-off between hosts. This
40 observation confirms that this classical theoretical prediction lacks empirical support. We
41 discuss the reasons for this discrepancy between theory and experiment in the light of our
42 results. The analysis of full genome consensus sequences of the evolved lineages established
43 that all mutations shared between lineages were host-specific. A low degree of parallel
44 evolution was observed, possibly reflecting the various adaptive pathways available for TEV
45 in each host. Altogether, these results reveal a strong adaptive potential of TEV to new hosts
46 without severe evolutionary constraints.

47

48 **Introduction**

49 The environment of a pathogen, and consequently the source of selection pressures is,
50 for the majority of its life-cycle, its host. The particular host environment presents different
51 forms of heterogeneity, even on short evolutionary time scales: from different cell types
52 within a tissue, to different species (Thomas et al. 2002). Multihost pathogens, which
53 encounter the latter, more extreme case of heterogeneity, should be able to cope with different
54 biotic conditions but also, due to their parasitic reproduction strategy, exploit different cellular
55 mechanisms and resources for their own growth and reproduction. The adaptation to various
56 hosts is supposed to be constrained by the fact that mutations advantageous in one
57 environment tend to be disadvantageous in another environment (Gandon 2004).

58 Simultaneous adaptation to various environments has important implications both for
59 applied and basic research. In recent years, emerging diseases have represented an important
60 threat for public health and agriculture (Anderson et al. 2004; Woolhouse, Haydon, Antia
61 2005; Parrish et al. 2008). In many instances, an emerging disease appears when a pathogen
62 jumps to another host it was not infecting (or no longer infecting) and adapts (or re-adapts) to
63 it. The increasing threat imposed by emerging and re-emerging diseases should prompt the
64 improvement of predictive models and the design of preventive strategies to control emerging
65 disease outbreaks. To do so, we still need to understand the mechanisms that make pathogen
66 populations able to spillover from their reservoirs, successfully infect a new naïve host, adapt
67 to it, and spread among the population of the new host.

68 In terms of basic research, simultaneous adaptation to various environments has been
69 an important topic in evolutionary biology for many years. It is thought to be one powerful
70 mechanism for the maintenance of genetic diversity (Kawecki 1994) and to play a crucial role
71 in triggering the divergence of incipient species (Via 1990). Indeed, one way a pathogen can
72 get around the evolutionary constraints imposed by the multiplicity of hosts is specialization

73 (or local adaptation) to one host species or to a small group of host species within the host
74 range. Host specialization can be the first step towards the evolution of host races and further
75 towards speciation (e.g. Filchak, Roethele, Feder 2000). Local adaptation to one host is
76 predicted to have a cost, manifested through worse performance of the locally adapted
77 pathogen on a non-native host (Kawecki, Ebert 2004). A recent literature survey established
78 that, when looked for, a significant pattern of local adaptation is usually found while the cost
79 of adaptation is, at best, small and in many cases non-significant (Hereford 2009).

80 Taking into account the potential for local adaptation, for any multihost pathogen
81 specific lineages are likely to fall along a gradient, which in principle ranges from absolute
82 specialists, those pathogens only able to infect and reproduce in a single host, to absolute
83 generalists, able to infect and reproduce equally well within any species of the host range. The
84 position of a particular pathogen lineage along this gradient depends on: i) environmental
85 factors, ii) the genetic diversity within each lineage and iii) the global fitness of each
86 evolutionary strategy. First, the environmental factors affecting the evolutionary fate of a
87 multihost pathogen are the frequency and distribution of each species of the host range, which
88 strongly influence the migration rate between host species (Woolhouse, Taylor, Haydon 2001;
89 Ravigne, Dieckmann, Olivieri 2009). For example, for a plant pathogen a monoculture field is
90 clearly a situation favouring specialist genotypes, while a diverse ecosystem is expected to
91 favour generalist genotypes. An additional environmental factor for vector-transmitted
92 pathogens is the availability and the specialization of the vectors. Indeed, vectors usually feed
93 on hosts and themselves have a preference towards one or more hosts, thus strongly
94 influencing the probability of transmitting the parasite from one species to another
95 (Woolhouse, Taylor, Haydon 2001). Second, the genetic diversity of a pathogen population
96 within one host species conditions the chances of appearance and selection of mutations that
97 allow for a sufficient infectivity and reproduction rate to maintain the pathogen within another

98 host species. The genetic diversity itself depends on population genetic parameters such as the
99 mutation rate, the reproduction rate, the intra-host effective population size, and population
100 structure. Third, generalist evolutionary strategies have been predicted to be disadvantaged
101 for three distinct reasons: (1) antagonistic pleiotropy, *i.e.* the fact that mutational effects are
102 negatively correlated between hosts, is thought to play a preponderant role in many cases (Via
103 1990; Fry 1996; Gandon 2004). (2) Since generalists alternate between hosts, some pathogen
104 genes can be expressed during infection in one host and not expressed in another host. The
105 absence of expression part of the time implies the absence of selection acting on these loci
106 and the potential for accumulation of mutations, which will be deleterious when expressed in
107 the other host (Kawecki 1994). (3) The alternation between hosts by generalists implies that
108 selection pressures specific to a particular host act only part of the time, slowing down the
109 rate of allele fixation and elimination compared to that of a specialist (Whitlock 1996). Field
110 studies have shown the disadvantage of generalists in a diversity of systems (e.g. Poulin 1998,
111 fish parasites; Poulin, Mouillot 2004, helminth parasites of birds; Malpica et al. 2006, viruses
112 and weeds; Straub, Ives, Gratton 2011, endoparasitoids and aphids). In other systems,
113 however, the generalist strategy seems to be favoured (e.g. Krasnov et al. 2004, ectoparasitic
114 fleas and small mammals; Hellgren, Pérez-Tris, Bensch 2009, *Plasmodium* sp. and birds).
115 This discrepancy between theoretical predictions and some field data is likely to be due to the
116 influence of other factors (environmental and genetic diversity) and to the lack of universality
117 of the costs of adaptation (Hereford 2009).

118 An efficient way to investigate the factors and mechanisms of specialization and the
119 implications for the evolution of specialist and generalist strategies is to use experimental
120 evolution (Kawecki, Ebert 2004). It allows explicitly addressing hypotheses about the role of
121 particular ecological and genetic factors that promote or hinder local adaptation, maintaining
122 other factors fixed. A number of experimental evolution studies on specialization and

123 generalist/specialist strategies have already been published using a variety of experimental
124 systems (e.g. Turner, Elena 2000, *Vesicular stomatitis virus* (VSV)/hamster, human or canine
125 cell; Magalhaes et al. 2009, spider mites/host plants; Legros, Koella 2010,
126 microsporidia/mosquito). Almost all experimental evolution studies show a pattern of host
127 specialization (Novella et al. 1999; Turner, Elena 2000; Cooper, Scott 2001; Greene et al.
128 2005; Wallis et al. 2007; Agudelo-Romero, de la Iglesia, Elena 2008; Coffey et al. 2008;
129 Magalhaes et al. 2009; Vasilakis et al. 2009; Legros, Koella 2010; Coffey, Vignuzzi 2011;
130 Deardorff et al. 2011), but only in some cases is specialization accompanied by a cost
131 (Weaver et al. 1999; Turner, Elena 2000; Wallis et al. 2007; Agudelo-Romero, de la Iglesia,
132 Elena 2008; Vasilakis et al. 2009; Legros, Koella 2010; Deardorff et al. 2011). Finally,
133 evidence for the cost of being a generalist are even less frequent (Coffey et al. 2008; Legros,
134 Koella 2010).

135 Among the existing experimental evolution studies, those using RNA viruses are
136 overrepresented. There are various reasons for this overrepresentation. First, RNA viruses are
137 likely to evolve in a reasonable experimental time because they have large population sizes,
138 fast replication rates, and a high mutation rate due to the absence of proofreading activity of
139 the RNA polymerase (Elena, Sanjuan 2007). Second, the small genome size allows access the
140 full genome sequence of the evolved lineages and to reveal the genetic changes that underlie
141 phenotypic evolution. Third, RNA viruses present a large diversity of evolutionary strategies.
142 Examples in animals range from highly specialized viruses, such as the *Poliovirus*
143 (*Picornaviridae*) which only infects humans, to the *Influenza A virus* (*Orthomyxoviridae*),
144 able to infect hosts from distant phylogenetic groups (Wolfe, Dunavan, Diamond 2007).
145 Diversity is also found in RNA plant viruses, with some viruses such as the *Tobacco etch*
146 *potyvirus* (TEV), which is a relative specialist compared to the *Cucumber mosaic*
147 *cucumovirus*, known to infect members of 16 different plant families. In this context, a group

148 of RNA viruses that has attracted a lot of attention is the arboviruses (arthropod-born viruses),
149 whose infective cycle includes obligate alternation between vertebrate and insect hosts. The
150 constraints imposed by the evolutionary trade-offs have been proposed as a likely cause of
151 their relative genetic homogeneity and constancy over time (Woelk, Holmes 2002).

152 It is important to note that the majority of these experimental evolution studies use cell
153 cultures as “hosts”. Cell cultures are simple, homogeneous environments, with no tissue
154 structure and no immune pressure, and probably represent conditions where selection
155 pressures are constant and unidirectional. Actually, it has been shown that local adaptation
156 obtained by evolution in mammalian cell cultures cannot be extrapolated to the corresponding
157 organism (Ciota et al. 2007). Moreover, experimental evolution conducted *in vivo* and in the
158 corresponding cell culture lead to very different results (Coffey et al. 2008). Additionally, all
159 the studies cited above compare the outcome of one evolutionary history alternating between
160 two hosts with the two corresponding single-host evolutionary histories. It is thus possible
161 that the results in the alternate evolutionary history are not due to alternation of host
162 environment, but are idiosyncratic to this specific treatment. In other words, when technically
163 and experimentally possible, it is desirable to have, on top of the technical “within
164 evolutionary history” replication, different generalist and specialist evolutionary histories.
165 This extra level of replication allows for more robust conclusions on the existence of a
166 specialist-generalist trade-off for a particular pathogen. Among the published evolution
167 experiments with viruses, only four were *in vivo* experiments (Wallis et al. 2007; Agudelo-
168 Romero, de la Iglesia, Elena 2008; Coffey et al. 2008; Deardorff et al. 2011). Three of these
169 studies were designed to evaluate the evolutionary constraints imposed by vector transmission
170 on adaptation (Wallis et al. 2007; Coffey et al. 2008; Deardorff et al. 2011). The other study
171 used only two hosts and did not contain any evolutionary treatment with host alternation
172 (Agudelo-Romero, de la Iglesia, Elena 2008).

173 For the present study, we designed an evolution experiment to study local adaptation
174 of TEV to four hosts within its natural host range. The hosts were complete plants, with all
175 the complexity this completeness implies for the dynamics of viral infection. Moreover, we
176 put a special emphasis on the replication of specialist and generalist evolutionary histories:
177 four “single host” and three “alternate host” evolutionary histories were experimentally
178 derived with a high “within history” replication level. With this experimental set up, we
179 aimed at collecting data in conditions that were “as natural as possible” and would allow for
180 drawing robust and broad conclusions on generalist and specialist strategies, as well as on the
181 conditions for local adaptation. The infectivity and virulence characteristics of the
182 experimentally evolved lineages were analysed and their full genome consensus sequences
183 obtained. It has to be noted here that testing for local adaptation and the cost of adaptation
184 requires measuring fitness or variables tightly associated to it. Here, the ideal approach would
185 have been to evaluate the viral population growth rate, but this approach was hardly
186 experimentally tractable and we decided to use infectivity and virulence as fitness proxies
187 instead. Infectivity is usually a good variable to evaluate local adaptation because it is a clear
188 component of pathogen fitness and it is monotonically related to it (Kawecki, Ebert 2004). On
189 the other hand, virulence is usually under stabilizing selection, with intermediate levels of
190 virulence corresponding to the fitness optimum (Frank 1993; Jensen et al. 2006; Fraser et al.
191 2007; de Roode, Yates, Altizer 2008). The existence of this fitness optimum potentially
192 implies a non-monotonic relationship between virulence and fitness, depending on the portion
193 of the selection gradient where the measures are taken. Moreover, it is possible that the virus
194 has a different virulence optimum in the four hosts used, making the relation between
195 virulence and fitness even more complex. Virulence is thus less suitable for evaluating local
196 adaptation, although it is often a useful parameter to understand the implications of the
197 evolutionary and coevolutionary processes in host-pathogen systems.

198

199 **Material and methods**

200

201 *Virus and plants*

202 Our model system is *Tobacco etch potyvirus* (TEV). TEV is a member of the
203 *Potyvirus* genus within the *Potyviridae* family and has a moderately wide host range (Shukla,
204 Ward, Brunt 1994). It has a positive sense single-strand RNA genome of 9.5 kb that encodes a
205 large polyprotein, which is auto-catalytically cleaved into ten multifunctional mature viral
206 proteins (Riechmann, Lain, García 1992). Recently, an overlapping ORF coding a small
207 additional protein after frameshifting has been discovered (Chung et al. 2008). The genome
208 replication in *Potyviridae* is performed by a virus-encoded RNA-dependent RNA-polymerase
209 that lacks proofreading activity. TEV mutation rate is thus high, estimated to be around 10^{-5} to
210 10^{-6} mutations per site and per generation (Tromas, Elena 2010).

211 Plasmid pMTEV contains the TEV genome (Bedoya, Daròs 2010), and was a
212 generous gift by Dr. J.A. Daròs. The TEV genome used to generate this clone has been
213 isolated from *Nicotiana tabacum* (Carrington et al. 1993) and its sequence is published
214 elsewhere (Carrasco et al. 2007). A stock of infected tissue was generated before starting the
215 evolution experiment (see Supplemental material for details).

216 Four host species were used: *N. tabacum*, *Nicotiana benthamiana*, *Datura stramonium*
217 and *Capsicum annuum*. They all belong to the *Solanacea* family and TEV produces systemic
218 symptoms in all of them. For all the experimental steps, plants were maintained in a
219 greenhouse at 25 °C and a 16 h photoperiod.

220

221 *Viral accumulation in each host*

222 Before starting the experimental evolution, viral accumulation in each of the four hosts
223 was measured by RT-qPCR (see Supplemental material for details) to ensure that
224 transmission took place when the viral load had reached a plateau and that the quantity of

225 virus used for transmission was equivalent for the four host species. This was done to equalize
226 the size of the transmission bottleneck in every experimental evolution lineage, and thus
227 differences between lineages could not be attributed to differential influence of genetic drift.

228 The obtained accumulation curve indicated that at 7 dpi, viral accumulation had
229 reached a plateau and that viral genomes accumulated at the same level in *N. tabacum*, *N.*
230 *benthamiana* and *D. stramonium* (around 2.2×10^7 viral RNA molecules per 100 ng of total
231 RNA) and at a lower level in *C. annuum* (3.2×10^6 viral RNA molecules per 100 ng of total
232 RNA).

233

234 *Experimental evolution*

235 The experimental evolution design contained seven evolutionary histories (figure 1).
236 In four of them, the viruses were serially passaged in the same host, hereafter denominated as
237 lineages Nb (*N. benthamiana*), Ds (*D. stramonium*), Ca (*C. annuum*) and Nt (*N. tabacum*). In
238 the three other lineages, the viruses were serially passaged on alternate hosts using the
239 following pairs: (*N. benthamiana*, *N. tabacum*– hereafter NbNt), (*N. tabacum*, *C. annuum* –
240 NtCa) and (*D. stramonium*, *C. annuum* – DsCa). The first type of evolutionary history
241 represents the conditions for the selection of a specialist strategy, whereas the second type is
242 an experimental approximation of conditions selecting for a more generalist strategy. Each
243 evolutionary history was replicated ten times.

244 To initiate the experimental evolution, a sap was prepared with 300 mg of infected
245 tissue (from the previously described stock) and 450 μ L of inoculation buffer. For each
246 replicate, two plants were mechanically inoculated with 5 μ L of this sap on one leaf. All
247 lineages were thus started with a genetically homogeneous viral population and *de novo*
248 mutation was the only source of raw material on which selection and genetic drift could act
249 (Elena, Lenski 2003). For the subsequent passages, at 7 dpi, the aerial part of one of the two

250 plants in each lineage was collected. If the two plants presented symptoms, the plant to collect
251 was chosen randomly. If only one presented symptoms, this one was collected. In both cases,
252 the inoculated leaf was removed and a sap was prepared with 300 mg of symptomatic leaf
253 tissue in 400 μ L of inoculation buffer if the infected tissue was from *C. annuum*, and 150 mg
254 of infected tissue in 1 mL of inoculation buffer for the other host species, so that each
255 infection was started with similar amounts of viral RNA. For each lineage, two plants were
256 then inoculated on one leaf with 5 μ L of sap. Fifteen serial passages were performed. One of
257 the Ca lineages was lost during experimental evolution.

258 The experimental procedure removes the natural vector of TEV (aphids) and thus
259 represents a simplified version of the virus life-cycle. This choice was made because using
260 aphid transmission would have greatly reduced the number of plants we could have infected
261 and consequently the replication level of and within evolutionary history. The evolutionary
262 implications of this choice are discussed below.

263

264 *Infectivity and virulence measurement*

265 After the 15th passage, infected tissue was collected from each lineage and the viral
266 RNA content measured by RT-qPCR (see Materials and Methods). The obtained
267 quantification was used to prepare saps of equal viral RNA concentration. Each of these saps
268 was mechanically inoculated (5 μ L of sap on one leaf) on three plants of each of the four host
269 species. This way, we had a complete crossed design, fully replicated three times. For
270 practical reasons, the inoculation was spread over four days, with the replicate lineages within
271 an evolutionary history split between the days. Additionally, each day three plants of each
272 species were inoculated with inoculation buffer as non-virus controls, and another three plants
273 were inoculated with a sap (at the same concentration of viral RNA as the evolved lineages)
274 made with the TEV stock used to start the experiment, and representing the ancestor for all

275 the evolved lineages. Before inoculation, the aerial part of each plant was measured (from the
276 basis of the stem to the apex) with a precision of 0.5 cm. At 21 dpi, each plant was checked
277 individually and the presence of symptoms was noted, to then calculate the infectivity. The
278 aerial part was measured with a precision of 0.5 cm and weighted (with a Kern 440-35N
279 balance, Kern and Sohn GmbH) with a precision of 10 mg. We define virulence as the degree
280 of damage caused to a plant by viral infection, and it is negatively correlated with host fitness
281 (Shaner et al. 1992; Sacristán, García-Arenal 2008). We calculated the virulence expressed in
282 size as:

$$283 \quad \text{Vir}_{size}(E_i H_j) = 1 - \Delta size(E_i H_j) / \Delta size(control)$$

284 where $\text{Vir}_{size}(E_i H_j)$ is the virulence expressed on size of the i^{th} replicate of evolutionary history
285 E when inoculated on the j^{th} replicate of host H and $\Delta size$ is the difference in size between the
286 day of infection and 21 dpi. A similar virulence index was obtained from the weight,
287 $\text{Vir}_{weight}(E_i H_j)$. However, $\Delta weight$ cannot be calculated directly because it is impossible to
288 weigh the plant before inoculation. We thus established the correlation between weight and
289 size for each species for plants of the same age as the ones we inoculated using an
290 independent cohort of healthy plants reared in the same conditions as the one used for
291 infectivity and virulence measurements. Using the correlation for each species and the size at
292 inoculation, we could estimate the weight at inoculation for each plant and thus estimate
293 $\Delta weight$.

294

295 *Genomic consensus sequence determination in the evolved lineages*

296 Total RNA was extracted from infected tissue of the 69 experimentally evolved
297 lineages. The TEV genome was amplified in three overlapping fragments and the sequencing
298 of the amplification products was outsourced (see supplemental material for details). We
299 obtained the consensus sequence from nucleotide 48 to nucleotide 9492, *i.e.* 99% of the full

300 genome and 100% of the coding sequence. The average coverage was 2.41. The genomes
301 were assembled and the mutations were identified using the Staden 2.0.0b7 package.

302 Our estimates of virulence and infectivity were done at the population level and we did
303 not explore the variability of these variables within each replicate lineage. For this reason, we
304 also directly sequenced PCR amplified virus population cDNAs rather than sequencing
305 multiple clones isolated from the population. This consensus sequencing approach allows for
306 detection of the dominant nucleotide at each base position. When multiple clean sequencing
307 reads showed clearly the presence of two peaks at one position, the lineage was considered to
308 be polymorphic at that position. However, it is impossible to measure the frequency of each
309 allele with this method and mutations could be present in the virus population at frequencies
310 lower than the detection threshold for a chromatogram. The real within-population diversity is
311 thus higher than the one reported here.

312

313 **Results**

314 *Infectivity and virulence*

315 The analysis of infectivity and virulence data was performed on a data set containing
316 only the data from lineages that had at least one mutation compared to the initial sequence of
317 the ancestral TEV infectious clone. This reduced the data set from 69 to 53 independent
318 lineages (6 to 9 independent lineages per evolutionary history).

319 A nominal logistic model with evolutionary history, host plant and their interaction
320 was built to analyse their effect on infectivity. The “evolutionary history replication” was
321 nested within the “evolutionary history” factor. This analysis revealed that all factors had a
322 significant effect on infectivity (table 1, figure 2). Figure 2 also shows the behaviour of the
323 ancestral virus, not included in the previously described nominal logistic model. The
324 infectivity of the ancestral virus is in the mid-low range of infectivities of the derived

325 lineages, except on *N. tabacum* where it has an infectivity of one. An ANOVA with the same
326 factors as above was performed with the virulence expressed on size and on weight as
327 variables. In this case, the replication within evolutionary history was taken as a random
328 factor. The method used was restricted maximum likelihood. For the two indices of virulence,
329 evolutionary history did not have a significant effect whereas host and the interaction between
330 host and evolutionary history had a significant effect (table 2, figure S1). The significant
331 interaction for the three considered variables indicates that the different evolutionary histories
332 produced distinct infection characteristic in the four hosts. Figure S1 also shows the virulence
333 pattern of the ancestral virus across the four hosts (data not included in the previously
334 described ANOVAs): its virulence is in the mid-high range of those of the derived lineages.

335 Using the full data set of the derived lineages, we could also ask more specific
336 questions: do we have a signature of host specialization? Does this depend on the
337 specialist/generalist characteristic of the evolutionary history? To answer these questions,
338 each evolutionary history was classified as “specialist” (one host) or as “generalist” (two
339 hosts in alternation). Additionally, each inoculation was classified as “local” or “foreign”: a
340 local inoculation is an inoculation on a host present during the experimental evolution for this
341 precise lineage (for example, NbNt is local on *N. benthamiana* and *N. tabacum*; and Nb is
342 local on *N. benthamiana*), whereas a foreign inoculation is an inoculation on a host which was
343 not present during the experimental evolution (for example, NbNt is foreign on *D.*
344 *stramonium* and *C. annuum*; and Nb is foreign on *N. tabacum*, *D. stramonium* and *C.*
345 *annuum*). A nominal logistic model was constructed with “host”, “specialist/generalist” and
346 “local/foreign” and their double and triple interactions as factors to determine the effect on
347 infectivity. The “specialist/generalist” and all interactions containing it did not have any
348 significant effect. “Host”, “local/foreign” and their interaction had a significant effect (table
349 3) and the interaction was due to the fact that on *D. stramonium*, the local lineages had an

350 infectivity that was two times larger than the foreign ones (figure 3a). This effect was present,
351 to a much lower extent, on *N. tabacum* while on the two other hosts local and foreign lineages
352 had identical infectivity. Therefore, this result indicates a strong local adaptation of lineages
353 that have evolved all or part of the time on *D. stramonium*. An ANOVA with the same factors
354 was then performed to analyse the influence on the two virulence indices. For virulence
355 expressed on size, the statistical results were very similar to those for infectivity (table 4), but
356 the origin of the significant interaction was different (figure 3b): virulence expressed on size
357 was identical for local and foreign lineages on all hosts except on *C. annuum*, where the local
358 lineages had a higher virulence than the foreign ones. For virulence expressed on weight, the
359 factors “host” and “local/foreign” had a significant effect and no other effect was significant.
360 The “local/foreign” effect was due to a higher virulence of local lineages, particularly on *D.*
361 *stramonium* and *C. annuum* (figure 3c).

362 Another way of looking at local adaptation, which also allows determining whether a
363 cost of adaptation exists, is to perform analyses by pairs of single host evolutionary histories.
364 Indeed, with the complete factorial design used, the three variables (infectivity, virulence
365 expressed on size, and virulence expressed on weight) were measured for each pair of single
366 host evolutionary histories in the two corresponding hosts and we could perform the classical
367 test for local adaptation (Kawecki, Ebert 2004). For each of the six pairs of single-host
368 evolutionary histories and for each host, the difference between the trait value of the local (or
369 native) evolutionary history and that of the foreign (or non-native) evolutionary history was
370 calculated. If this difference is positive, it indicates local adaptation, whereas if it is negative,
371 it indicates local maladaptation. This analysis assumes that our variables are positively related
372 to fitness, which is the case for infectivity and is arguably the case for virulence (see
373 Introduction and Discussion). For each variable, a distribution of twelve values was obtained
374 and it was possible to test whether the differences were on average positive. The mean is

375 positive for the three variables but not significantly different from zero for any of them:
376 infectivity (1-tailed t -test: $p = 0.096$; Wilcoxon rank test: $p = 0.071$), virulence expressed on
377 size (1-tailed t -test: $p = 0.069$; Wilcoxon rank test: $p = 0.17$) and virulence expressed on
378 weight (1-tailed t -test: $p = 0.064$; Wilcoxon rank test: $p = 0.073$). However, the combined
379 Fisher test reveals an overall significant local adaptation (combining 1-tailed t -tests: $p =$
380 0.017 ; combining rank tests: $p = 0.029$). Then, using the same differences, it is possible to
381 evaluate the frequency at which the adaptation in one host comes to a cost of adaptation in the
382 other host. For the six pairs of single host evolutionary histories, we compared the differences
383 previously calculated in each of the two hosts. For example, we compared the (local –
384 foreign) difference in infectivity of the Nb and Nt lineages when inoculated on *N.*
385 *benthamiana* with their difference in infectivity when inoculated on *N. tabacum*. If the two
386 differences are positive, then we have a case of local adaptation with a cost of adaptation. Out
387 of the 18 cases (six pairs \times three variables), there are seven cases of local adaptation with a
388 cost of adaptation, ten cases of local adaptation of one of the evolutionary histories without
389 cost expressed in the other host and one case of double maladaptation. Out of the ten cases of
390 adaptation without cost, five concern lineages evolved on *D. stramonium*: the adaptation to *D.*
391 *stramonium* seems thus to be compatible with the infection and virulence expression in other
392 hosts. Finally, for three of the six pairs of single host evolutionary histories, data are available
393 for the corresponding alternate host evolutionary histories. Out of the seven cases where a
394 cost of adaptation has been identified, we have data for the corresponding alternate host
395 evolutionary history in four cases. There is no clear cost of being a generalist (which would
396 be expressed as a lower value for the generalist strategy than for the locally adapted specialist
397 strategy in each host) in any of these four cases, confirming the idea that the specialist-
398 generalist trade-off is not present in this experimental system. Figure 4 represents the data for
399 virulence expressed on weight for the (Nt, Ca, NtCa) group of evolutionary histories, as an

400 example of a situation with local adaptation, a cost of adaptation and no cost for the
401 corresponding generalist.

402
403 *Consensus sequences*

404 A total of 107 independent mutations occurring at 91 different loci were identified
405 with a range of zero to six mutations per independent evolved lineage in the 69 independently
406 evolved lineages (see figure 5 for a graphical representation and table S1 for a complete list).
407 The transition:transversion ratio was 7.2. Sixty-two mutations were synonymous and 45 non-
408 synonymous, the biological interpretation hereof is discussed below. Out of the 107 mutations
409 observed, 26 (24%) were not unique and out of the 91 polymorphic loci identified, 10 (11%)
410 were affected in multiple independent lineages. All the non-unique mutations were repeated
411 in lineages that shared a host in their evolutionary history, except in one case (A5409G
412 present in one Nb and one NtCa lines). The very large majority (24 out of 26) of non-unique
413 mutations were thus specific to one of the hosts used in the experiment.

414 We then tested whether the regions encoding for each of the eleven mature proteins
415 had a different rate of mutation accumulation and whether this depended on the evolutionary
416 history of the lineage. To do this, we calculated the number of mutations per nucleotide site
417 for each evolutionary history and for each mature protein. A Scheirer-Ray-Hare
418 nonparametric two-way ANOVA (Sokal, Rohlf 1995) showed that the mutation accumulation
419 per nucleotide is not globally different from one evolutionary history to another, but that
420 mutations are not randomly scattered across mature proteins and that this distribution is
421 different from one evolutionary history to another (Table 5). Namely, there is an
422 overrepresentation of mutations in P1 in lineages Ca, Nt and NtCa, of mutations in HC-Pro in
423 lineages Ds and DsCa, of mutations in P3 in lineages Nb and NbNt, of mutations in CI in
424 lineages NtCa, of mutations in 6K2 in lineages Nb, of mutations in VPg in lineages Ds and of
425 mutations in NIa-Pro in lineages Ca.

426 Focussing on individual consensus sequences, it is possible to identify patterns of
427 mutations that suggest a potential relationship between genotype and phenotype. First, among
428 the Nb lineages, there are two mutations repeated three times, and there is no case where both
429 are present in the same consensus sequence. Moreover, one of them (A3013G, non-
430 synonymous K→E) was also found in three independent NbNt lineages whereas the other one
431 (A5551G, non-synonymous K→E) was not found in any of the NbNt sequence. This suggests
432 that there are at least two mutations that can improve the adaptation to infection and
433 reproduction in *N. benthamiana* and that one (A3013G) is neutral or advantageous whereas
434 the other one (A5551G) is deleterious in *N. tabacum*. Second, an identical combination of
435 four mutations (A3013G, C3816A, A6805G, and G8169A) has been found in lineages Nb8
436 and NbNt6. Except the mutation A3013G, the other three have not been found in any other
437 lineage besides these two. This could be due to the existence of epistatic interactions in this
438 group of mutations. Third, lineages Ca1 and Ca2 have lost the ability to produce systemic
439 infections when inoculated on *N. tabacum* and on *D. stramonium* and form only local lesions
440 on the inoculated leaf, but still produce systemic symptoms on *C. annuum* and *N.*
441 *benthamiana*. The analysis of their sequence revealed only one synonymous mutation in Ca1
442 (C4425U in the CI region) and two synonymous mutations in Ca2 (A183G and U8574C in
443 the P1 and CP region, respectively).

444 The relationship between the mutations and the phenotype observed are clearly
445 correlational and speculative for now and the evolutionary mechanisms suggested by the
446 analysis of the sequences have to and will be directly tested in future site-directed
447 mutagenesis experiments.

448

449 **Discussion**

450 TEV evolved differently when experimentally exposed to different evolutionary histories, as
451 evidenced by the phenotypic traits measured, as well as by the sequence changes observed in
452 the evolved lineages. The ancestral virus and the lineages evolved on *N. tabacum*, the
453 ancestral host, present different patterns of infectivity and virulence. This is likely due to an
454 adaptation to the specific conditions of the experimental evolution and in particular to the
455 mechanical transmission, which totally removes the selection due to the aphid transmission.
456 The transmission mode is actually known to affect viral evolution both at the phenotypic and
457 at the genotypic levels (Wallis et al. 2007; Jerzak et al. 2008). Due to these differences
458 between the ancestor and the Nt lineages, we focused the analysis and the discussion on the
459 comparison of the experimental evolutionary histories between them.

460 We observe that the viral populations have adapted to the hosts that they encountered
461 during their specific evolutionary histories. The analyses performed on pairs of populations
462 further confirm the pattern of local adaptation. In some systems, local adaptation comes at a
463 cost (Hereford 2009), expressed as a worse performance of the locally adapted populations
464 when measured in a foreign environment. In the case of TEV on the four hosts used, the cost
465 of adaptation is not a general feature, as it appears only in some cases. The results for
466 infectivity and virulence were congruent in direction, albeit not in magnitude for all the
467 analyses performed.

468 The consensus sequences of the evolved populations revealed common mutations
469 between lineages sharing host species in their evolutionary histories. Such parallel evolution
470 is usually interpreted as the fixation of a mutation with a beneficial effect (Wood, Burke,
471 Rieseberg 2005), and in this precise case, the host-specific mutations likely increase the level
472 of adaptation to the particular host species. One potential concern about the experimental
473 protocol is that the removal of the aphid vector from the virus life-cycle would relax selection

474 pressures at some loci, e.g. in the N-terminal part of HC-Pro known to be involved in aphid
475 transmission (Blanc et al. 1998) and that part of the mutations observed would be due to this.
476 The absence of common mutations between lineages that do not share a host excludes the
477 relaxed selection pressure of aphid transmission as the main evolutionary force shaping
478 sequence during our evolution experiment.

479 Overall, the level of parallel evolution obtained here is below those in other studies
480 comparing genomes from lineages sharing evolutionary history (e.g., Bull et al. 1997;
481 Martínez-Picado et al. 2000; Wichman, Millstein, Bull 2005; Remold, Rambaut, Turner
482 2008). In particular, the experiment conducted by Remold et al. (2008) presents a number of
483 similarities with the present one: it involves experimental evolution of a RNA virus (VSV) by
484 serial transfers either in one of two cell cultures (human or hamster cells) or in the alternation
485 of the two cell types, followed by full genome sequencing of twelve evolved populations. The
486 authors found that 78% of the observed mutations were not unique and 55% of the
487 polymorphic loci were affected in multiple independent populations. The reasons for the
488 lower level of parallel evolution in our system are likely manifold. First, this level strongly
489 depends on the genetic architecture of adaptation, *i.e.* on the distribution of mutation effects in
490 the organisms and in the environments considered (Chevin, Martin, Lenormand 2010) but, the
491 shape of these distributions is rarely known and consequently the role of this factor in
492 determining the level of parallel evolution cannot be evaluated on concrete evolution
493 experiments. Second, the lower level of parallel evolution in our study could be due to the
494 narrower and less controlled bottlenecks between passages: we controlled the amount of viral
495 RNA inoculated but the number of virions actually starting the new infection can vary
496 considerably (Zwart, Daròs, Elena 2011). In any case, the number of virions starting a new
497 infection represents an extremely small fraction of those produced during the colonisation of
498 the host in the previous passage and a much smaller fraction than in the case of serial transfer

499 of virus in cell culture. This reduction in effective population size increases the role of genetic
500 drift and reduces the likelihood of fixation of the same mutation in two replicates of the same
501 evolutionary treatment. Third, populations of plant viruses are highly structured because (i)
502 infection within a leaf progresses only from one cell to adjacent ones, (ii) viral exclusion
503 phenomenon are frequent (Dietrich, Maiss 2003; Zwart, Daròs, Elena 2011) and reduce the
504 generation of diversity by recombination, and (iii) the order of colonization of the leaves is
505 fixed by the phloem distribution. All these factors generate much more opportunities for local
506 competition and exclusion of alleles, more heterogeneity in selection pressures and result in a
507 much smaller effective population size than a well-mixed cell culture. Finally, the low level of
508 parallel evolution might come from the fact that we performed a passage every week during
509 experimental evolution. One week might represent an insufficient time for the virus to
510 generate genetic diversity and for the selection to act on this diversity. However, first for
511 technical reasons and second owing to the drastic effect of TEV on some of the hosts used in
512 this experiment (see virulence levels on *N. benthamiana*), it was unrealistic to perform longer
513 passages.

514 For the analysis of consensus sequences, we decided not to interpret the
515 synonymous/nonsynonymous characteristic of the mutations as neutral/non-neutral.
516 Traditionally synonymous mutations have been considered neutral because they do not induce
517 amino acid changes and their proportion relative to the nonsynonymous ones has been used to
518 deduce the evolutionary forces at work. However, there is accumulating evidence that
519 synonymous mutations are not always neutral in DNA organisms (Kimchi-Sarfaty et al. 2007;
520 Amorós-Moya et al. 2010; Plotkin, Kudla 2011) and the equivalence of synonymous with
521 neutral is likely to be even weaker in RNA viruses, because tridimensional structure is a
522 determinant feature of RNA molecules and it is largely dependent on the sequence itself. The
523 RNA-based genome can thus itself potentially be a target of selection, because coding regions

524 serve additional functions other than determining the amino acid sequence in protein. These
525 additional functions include encapsidation or serving as a target of silencing (Cuevas,
526 Domingo-Calap, Sanjuán 2011). In TEV, it has been shown that certain synonymous
527 mutations had deleterious effect on the virus multiplication (Carrasco, de la Iglesia, Elena
528 2007) and the present study provides a striking example of non-neutrality of synonymous
529 mutations with the two lineages which lost the capacity to produce systemic infection on
530 certain hosts and only carry synomymous mutations.

531 Another interesting result of our study is the absence of cost of generalism, either
532 when we look at the full data set or when we can contrast the characteristics of one generalist
533 evolutionary history with the ones of the two corresponding specialists. This is probably
534 partly due to the fact that the cost of adaptation is not general and consequently, that the
535 adaptation to two hosts simultaneously does not represent a strong evolutionary constraint.
536 However, what might seem harder to explain is the pattern of reciprocal local adaptation with
537 a cost of adaptation and without any identified cost of generalism (e.g. figure 4). A pattern of
538 reciprocal local adaptation with a cost of adaptation can be due either to antagonistic
539 pleiotropy or to fixation of mutations beneficial in the local host and neutral in the other one.
540 Indeed, in potyviruses, we can exclude the accumulation of deleterious mutations at loci not
541 expressed in one of the hosts (Kawecki 1994) as the full genome is constantly expressed as a
542 single polyprotein. If the cost of adaptation was due to antagonistic pleiotropy, the generalists
543 should pay a cost. Antagonistic pleiotropy makes impossible the adaptation to two hosts at the
544 same time. If it was due to neutral mutations, they could perform as well as the best lineage in
545 each host. Our experimental results favour the second explanation. However, the previous
546 reasoning is based on phenotypic traits only and the sequence data indicate a low level of
547 parallel evolution, meaning that the adaptation is due to different mutations in replicate
548 lineages of an evolutionary history and between evolutionary histories sharing a host. This

549 weakens considerably the argumentation based on phenotypic traits and knowing really
550 whether the mutations obtained are antagonistically pleiotropic would require analysing them
551 one by one through site-directed mutagenesis of the wild-type TEV genome. Finally, another
552 potential, non-exclusive explanation for the absence of cost of generalism is that the
553 “generalist populations” could actually be a composite of two populations, each one being
554 adapted to one of the hosts. If this was the case, it should be visible on the consensus
555 sequence data with a highest proportion of identified polymorphic sites in generalist than in
556 specialist populations. This is however not the case: polymorphic sites represent 24.4% of the
557 sites where a change has been identified in generalist lineages compared to 28.8% in
558 specialist lineages (Fisher’s exact test, $p = 0.763$).

559 To sum up, our experimental evolution of TEV in a multihost context revealed a
560 pattern of local adaptation, characterized in particular by a higher infectivity on the host(s)
561 encountered during the experimental evolution. Local adaptation did not always come with a
562 cost of adaptation and the cost of generalists predicted by theoretical approaches was not
563 found. At the sequence level, host specific mutations were found but the general level of
564 parallel evolution was relatively low. Altogether, these experimental results reveal a strong
565 adaptation potential of TEV to new hosts without severe evolutionary constraints. Our
566 experiment, using an *in vivo* system and a larger replication of and within evolutionary
567 histories, points in the same direction as previous experimental evolution approaches on other
568 pathogens. The generality of the high potential of adaptation of pathogens to hosts, which are
569 new to them at the evolutionary scale considered, indicates that emerging diseases caused by
570 a diversity of organisms actually represents an important threat. The most efficient policies to
571 limit them are preventive and should target the steps prior to adaptation: limit the flow of
572 pathogens towards new host species and avoid generating evolutionary situations favourable
573 to adaptation to a new host.

574

575 **Supplementary material:** see supplementary files. Contains supplementary methods,
576 supplementary table 1 and supplementary figure 1.

577

578 **Acknowledgement:** We thank Francisca de la Iglesia and Àngels Pròsper for
579 excellent technical assistance and Mark Zwart and two anonymous reviewers for their helpful
580 comments on a previous version of the manuscript. This research was supported by the
581 Spanish Ministry of Science and Innovation grant BFU2009-06993 to SFE. SB was supported
582 by the JAE-doc program from CSIC and GL was supported by the Human Frontier Science
583 Program, grant RGP0012/2008.

584

585 **Literature cited**

586 Agudelo-Romero, P, F de la Iglesia, SF Elena. 2008. The pleiotropic cost of host-
587 specialization in *Tobacco etch potyvirus*. *Inf Genet Evol* 8:806-814.

588 Amorós-Moya, D, S Bedhomme, M Hermann, IG Bravo. 2010. Evolution in regulatory
589 regions rapidly compensates the cost of nonoptimal codon usage. *Mol Biol Evol*
590 27:2141-2151.

591 Anderson, PK, AA Cunningham, NG Patel, FJ Morales, PR Epstein, P Daszak. 2004.

592 Emerging infectious diseases of plants: pathogen pollution, climate change and
593 agrotechnology drivers. *Trends Ecol Evol* 19:535-544.

594 Bedoya, LC, J-A Daròs. 2010. Stability of *Tobacco etch virus* infectious clones in plasmid
595 vectors. *Virus Res* 149:234-240.

596 Blanc, S, ED Ammar, S Garcia-Lampasona, VV Dolja, C Llave, J Baker, TP Pirone. 1998.

597 Mutations in the potyvirus helper component protein: effects on interactions with
598 virions and aphid stylets. *J Gen Virol* 79:3119-3122.

599 Bull, JJ, MR Badgett, HA Wichman, JP Huelsenbeck, DM Hillis, A Gulati, C Ho, IJ
600 Molineux. 1997. Exceptional convergent evolution in a virus. *Genetics* 147:1497-
601 1507.

602 Carrasco, P, JA Daròs, P Agudelo-Romero, SF Elena. 2007. A real-time RT-PCR assay for
603 quantifying the fitness of *Tobacco etch virus* in competition experiments. *J Virol*
604 *Methods* 139:181-188.

605 Carrasco, P, F de la Iglesia, SF Elena. 2007. Distribution of fitness and virulence effects
606 caused by single-nucleotide substitutions in *Tobacco etch virus*. *J Virol* 81:12979-
607 12984.

608 Carrington, JC, R Haldeman, VV Dolja, MA RestrepoHartwig. 1993. Internal cleavage and
609 trans-proteolytic activities of the Vpg-Proteinase (NIa) of *Tobacco etch potyvirus in*
610 *vivo*. *J Virol* 67:6995-7000.

611 Chevin, L-M, G Martin, T Lenormand. 2010. Fisher's model and the genomics of adaptation:
612 restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution*
613 64:3213-3231.

614 Chung, BYW, WA Miller, JF Atkins, AE Firth. 2008. An overlapping essential gene in the
615 *Potyviridae*. *Proc Natl Acad Sci USA* 105:5897-5902.

616 Ciota, AT, AO Lovelace, SA Jones, A Payne, LD Kramer. 2007. Adaptation of two
617 flaviviruses results in differences in genetic heterogeneity and virus adaptability. *J*
618 *Gen Virol* 88:2398-2406.

619 Coffey, LL, N Vasilakis, AC Brault, AM Powers, F Tripet, SC Weaver. 2008. Arbovirus
620 evolution *in vivo* is constrained by host alternation. *Proc Natl Acad Sci USA*
621 105:6970-6975.

622 Coffey, LL, M Vignuzzi. 2011. Host Alternation of *Chikungunya virus* increases fitness while
623 restricting population diversity and adaptability to novel selective pressures. *J Virol*
624 85:1025-1035.

625 Cooper, LA, TW Scott. 2001. Differential evolution of *Eastern equine encephalitis virus*
626 populations in response to host cell type. *Genetics* 157:1403-1412.

627 Cuevas, JM, P Domingo-Calap, R Sanjuán. 2011. The fitness effects of synonymous
628 mutations in DNA and RNA viruses. *Mol Biol Evol* 28:in press.

629 Deardorff ER, KA Fitzpatrick, GVS Jerzak, PY Shi, L.D. Kramer, GD Ebel. 2011. *West Nile*
630 *virus* experimental evolution *in vivo* and the trade-off hypothesis. *PLoS Pathog*
631 7:e1002335.

632 de Roode, JC, AJ Yates, S Altizer. 2008. Virulence-transmission trade-offs and population
633 divergence in virulence in a naturally occurring butterfly parasite. *Proc Natl Acad Sci*
634 U S A 105:7489-7494.

635 Dietrich, C, E Maiss. 2003. Fluorescent labelling reveals spatial separation of potyvirus
636 populations in mixed infected *Nicotiana benthamiana* plants. *J Gen Virol* 84:2871-
637 2876.

638 Elena, SF, RE Lenski. 2003. Evolution experiments with microorganisms: The dynamics and
639 genetic bases of adaptation. *Nat Rev Genet* 4:457-469.

640 Elena, SF, R Sanjuan. 2007. Virus evolution: Insights from an experimental approach. *Annu*
641 *Rev Ecol Evol* 38:27-52.

642 Filchak, KE, JB Roethele, JL Feder. 2000. Natural selection and sympatric divergence in the
643 apple maggot *Rhagoletis pomonella*. *Nature* 407:739-742.

644 Frank, SA. 1993. Evolution of host-parasite diversity. *Evolution* 47:1721-1732.

645 Fraser, C, TD Hollingsworth, R Chapman, F de Wolf, WP Hanage. 2007. Variation in HIV-1
646 set-point viral load: epidemiological analysis and an evolutionary hypothesis. Proc
647 Natl Acad Sci U S A 104:17441-17446.

648 Fry, JD. 1996. The evolution of host specialization: Are trade-offs overrated? Am Nat
649 148:S84-S107.

650 Gandon, S. 2004. Evolution of multihost parasites. Evolution 58:455-469.

651 Greene, IP, EY Wang, ER Deardorff, R Milleron, E Domingo, SC Weaver. 2005. Effect of
652 alternating passage on adaptation of *Sindbis virus* to vertebrate and invertebrate cells.
653 J Virol 79:14253-14260.

654 Hellgren, O, J Pérez-Tris, S Bensch. 2009. A jack-of-all-trades and still a master of some:
655 prevalence and host range in avian malaria and related blood parasites. Ecology
656 90:2840-2849.

657 Hereford, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. Am Nat
658 173:579-588.

659 Jensen, KH, TJ Little, A Skorping, D Ebert. 2006. Empirical support for optimal virulence in
660 a castrating parasite. PLoS Biol 4:e197.

661 Jerzak, GVS, I Brown, PY Shi, LD Kramer, GD Ebel. 2008. Genetic diversity and purifying
662 selection in *West Nile virus* populations are maintained during host switching.
663 Virology 374:256-260.

664 Kawecki, TJ. 1994. Accumulation of deleterious mutations and the evolutionary cost of being
665 generalist. Am Nat 144:833-838.

666 Kawecki, TJ, D Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters 7:1225-
667 1241.

668 Kimchi-Sarfaty, C, JM Oh, IW Kim, ZE Sauna, AM Calcagno, SV Ambudkar, MM
669 Gottesman. 2007. A "silent" polymorphism in the MDR1 gene changes substrate
670 specificity. *Science* 315:525-528.

671 Krasnov, BR, R Poulin, GI Shenbrot, D Mouillot, IS Khokhlova. 2004. Ectoparasitic "jacks-
672 of-all-trades": Relationship between abundance and host specificity in fleas
673 (*Siphonaptera*) parasitic on small mammals. *Am Nat* 164:506-516.

674 Legros, M, JC Koella. 2010. Experimental evolution of specialization by a microsporidian
675 parasite. *BMC Evol Biol* 10:159.

676 Magalhaes, S, E Blanchet, M Egas, I Olivieri. 2009. Are adaptation costs necessary to build
677 up a local adaptation pattern? *BMC Evol Biol* 9:182.

678 Malpica, JM, S Sacristán, A Fraile, F García-Arenal. 2006. Association and host selectivity in
679 multi-host pathogens. *PLoS ONE* 1:e41.

680 Martínez-Picado, J, MP DePasquale, N Kartsonis, et al. 2000. Antiretroviral resistance during
681 successful therapy of HIV type 1 infection. *Proc Natl Acad Sci USA* 97:10948-10953.

682 Novella, IS, CL Hershey, C Escarmis, E Domingo, JJ Holland. 1999. Lack of evolutionary
683 stasis during alternating replication of an arbovirus in insect and mammalian cells. *J*
684 *Mol Biol* 287:459-465.

685 Parrish, CR, EC Holmes, DM Morens, EC Park, DS Burke, CH Calisher, CA Laughlin, LJ
686 Saif, P Daszak. 2008. Cross-species virus transmission and the emergence of new
687 epidemic diseases. *Microbiol Mol Biol R* 72:457-+.

688 Plotkin, JB, G Kudla. 2011. Synonymous but not the same: the causes and consequences of
689 codon bias. *Nat Rev Genet* 12:32-42.

690 Poulin, R. 1998. Large-scale patterns of host use by parasites of freshwater fishes. *Ecol Lett*
691 1:118-128.

692 Poulin, R, D Mouillot. 2004. The relationship between specialization and local abundance: the
693 case of helminth parasites of birds. *Oecologia* 140:372-378.

694 Ravigne, V, U Dieckmann, I Olivieri. 2009. Live where you thrive: joint evolution of habitat
695 choice and local adaptation facilitates specialization and promotes diversity. *Am Nat*
696 174:E141-169.

697 Remold, SK, A Rambaut, PE Turner. 2008. Evolutionary genomics of host adaptation in
698 *Vesicular stomatitis virus*. *Mol Biol Evol* 25:1138-1147.

699 Riechmann, JL, S Lain, JA García. 1992. Highlights and prospects of *Potyvirus* molecular
700 biology. *J Gen Virol* 73:1-16.

701 Sacristán, S, F García-Arenal. 2008. The evolution of virulence and pathogenicity in plant
702 pathogen populations. *Mol Plant Pathol* 9:369-384.

703 Shaner, G, EL Stromberg, GH Lacy, KR Barker, TP Pirone. 1992. Nomenclature and
704 concepts of pathogenicity and virulence. *Annu Rev Phytopathol* 30:47-66.

705 Shukla, DD, CW Ward, AA Brunt. 1994. *The Potyviridae*. Wallingford: CABI.

706 Sokal, RR, FJ Rohlf. 1995. *Biometry*. New York: W.H. Freeman and company.

707 Straub, CS, AR Ives, C Gratton. 2011. Evidence for a trade-off between host-range breadth
708 and host-use efficiency in aphid parasitoids. *Am Nat* 177:389-395.

709 Thomas, F, SP Brown, M Sukhdeo, F Renaud. 2002. Understanding parasite strategies: a
710 state-dependent approach? *Trends Parasitol* 18:387-390.

711 Tromas, N, SF Elena. 2010. The rate and spectrum of spontaneous mutations in a plant RNA
712 virus. *Genetics* 185:983-989.

713 Turner, PE, SF Elena. 2000. Cost of host radiation in an RNA virus. *Genetics* 156:1465-1470.

714 Vasilakis, N, ER Deardorff, JL Kenney, SL Rossi, KA Hanley, SC Weaver. 2009. Mosquitoes
715 put the brake on arbovirus evolution: experimental evolution reveals slower mutation
716 accumulation in mosquito than vertebrate cells. *PLoS Pathog* 5:e1000467.

717 Via, S. 1990. Ecological genetics and host adaptation in herbivorous insects - the
718 experimental study of evolution in natural and agricultural systems. *Annu Rev*
719 *Entomol* 35:421-446.

720 Wallis, CM, AL Stone, DJ Sherman, VD Damsteegt, FE Gildow, WL Schneider. 2007.
721 Adaptation of *Plum pox virus* to a herbaceous host (*Pisum sativum*) following serial
722 passages. *J Gen Virol* 88:2839-2845.

723 Weaver, SC, AC Brault, WL Kang, JJ Holland. 1999. Genetic and fitness changes
724 accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. *J Virol*
725 73:4316-4326.

726 Whitlock, MC. 1996. The red queen beats the jack-of-all-trades: The limitations on the
727 evolution of phenotypic plasticity and niche breadth. *Am Nat* 148:S65-S77.

728 Wichman, HA, J Millstein, JJ Bull. 2005. Adaptive molecular evolution for 13,000 phage
729 generations: A possible arms race. *Genetics* 170:19-31.

730 Woelk, CH, EC Holmes. 2002. Reduced positive selection in vector-borne RNA viruses. *Mol*
731 *Biol Evol* 19:2333-2336.

732 Wolfe, ND, CP Dunavan, J Diamond. 2007. Origins of major human infectious diseases.
733 *Nature* 447:279-283.

734 Wood, TE, JM Burke, LH Rieseberg. 2005. Parallel genotypic adaptation: when evolution
735 repeats itself. *Genetica* 123:157-170.

736 Woolhouse, MEJ, DT Haydon, R Antia. 2005. Emerging pathogens: the epidemiology and
737 evolution of species jumps. *Trends Ecol Evol* 20:238-244.

738 Woolhouse, MEJ, LH Taylor, DT Haydon. 2001. Population biology of multihost pathogens.
739 *Science* 292:1109-1112.

740 Zwart, MP, JA Daròs, SF Elena. 2011. One is enough: in vivo effective population size is
741 dose-dependent for a plant RNA virus. *PLoS Pathog* 7:e1002122.

742

743 **Tables.**

744 *Table 1.* Nominal logistic model on the infectivity data of the experimentally evolved

745 lineages.

746

<i>Factor</i>	<i>d.f.</i>	<i>Likelihood ratio χ^2</i>	<i>p</i>
Evolutionary history	6	115.92	<0.0001
Host	3	166.34	<0.0001
Evol. hist. \times host	18	83.01	<0.0001
replication [evol. hist.]	46	158.77	<0.0001

747

748

749 *Table 2.* ANOVA of the virulence data (expressed on size and on weight) of the
 750 experimentally evolved lineages.

751

Virulence expressed on size

<i>Factor</i>	<i>d.f. denominator</i>	<i>d.f. numerator</i>	<i>F ratio</i>	<i>p</i>
Evolutionary history	6	49.66	2.06	0.0752
Host	3	311.34	215.99	<0.0001
Evol. hist. × host	18	304.81	3.95	<0.0001
replication [evol. hist.] (Rdm)			<i>percentage of variance explained</i>	23.61%

Virulence expressed on weight

<i>Factor</i>	<i>d.f. denominator</i>	<i>d.f. numerator</i>	<i>F ratio</i>	<i>p</i>
Evolutionary history	6	43.23	0.805	0.5713
Host	3	317.78	10.937	<0.0001
Evol. hist. × host	18	311.51	1.843	0.0202
replication [evol. hist.] (Rdm)			<i>percentage of variance explained</i>	9.51%

752

753

754 *Table 3.* Effect of the type of evolutionary history (specialist or generalist) and of the type of
755 infection (foreign or local) on infectivity.

756

<i>Factor</i>	<i>d.f.</i>	<i>Likelihood ratio χ^2</i>	<i>p</i>
Host	3	150.52	<0.0001
Foreign/Local	1	7.58	0.0059
Specialist/Generalist	1	0.41	0.5222
Host \times F/L	3	24.35	<0.0001
Host \times S/G	3	2.04	0.5635
F/L \times S/G	1	0.71	0.4008
Host \times F/L \times S/G	3	2.53	0.4704

757

758

759 *Table 4.* Effect of the type of evolutionary history (specialist or generalist) and of the type of
 760 infection (foreign or local) on virulence.

761

Virulence expressed on size

<i>Factor</i>	<i>d.f.</i>	<i>F ratio</i>	<i>p</i>
Host	3	255.14	<0.0001
Foreign/Local	1	15.05	0.0001
Specialist/Generalist	1	0.03	0.8655
Host × F/L	3	5.62	0.0009
Host × S/G	3	1.20	0.3100
F/L × S/G	1	2.81	0.0943
Host × F/L × S/G	3	3.32	0.0199

Virulence expressed on weight

<i>Factor</i>	<i>d.f.</i>	<i>F ratio</i>	<i>p</i>
Host	3	17.590	<0.0001
Foreign/Local	1	6.460	0.0114
Specialist/Generalist	1	0.365	0.5459
Host × F/L	3	1.791	0.1482
Host × S/G	3	0.782	0.5043
F/L × S/G	1	0.727	0.3943
Host × F/L × S/G	3	1.009	0.3885

762

763

764 *Table 5.* Nonparametric analysis of variance for the number of nucleotide substitutions among
765 genes and for each experimental evolutionary history.

766

<i>Term</i>	<i>d.f.</i>	<i>H</i>	<i>P value</i>
Evolutionary history	6	10.81	0.0944
Genome region	9	30.80	0.0003
Evol. hist. × Gen. region	54	103.57	<0.0001

767

768

769 **Figure legends**

770

771 *Figure 1.* Overview of the experimental evolution procedure. The seven evolutionary histories
772 are represented: in continuous lines, the “specialist” history and in dotted lines, the
773 evolutionary histories with host alternation. Host species are codified as follows: Ds = *D.*
774 *stramonium*, Ca = *C. annuum*, Nt = *N. tabacum* and Nb = *N. benthamiana*.

775

776 *Figure 2.* Infectivity of the different evolutionary histories on the four experimental hosts.
777 Infectivity is expressed in terms of proportion of plants infected \pm 1 SEM.

778

779 *Figure 3.* Experimentally evolved local adaptation: infectivity (A), virulence expressed on
780 size (B) and virulence expressed on weight (C) of local (triangles) and foreign (circles) across
781 the four experimental hosts. All values are expressed \pm 1 SEM.

782

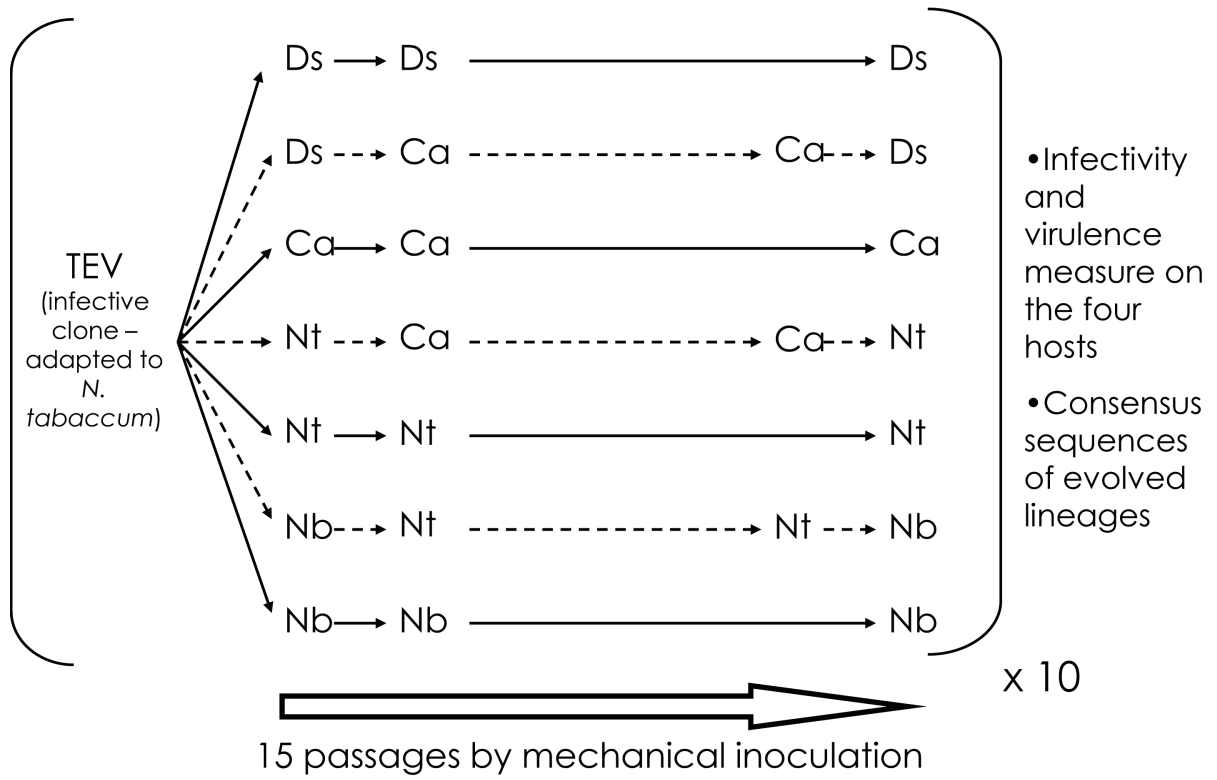
783 *Figure 4.* Virulence expressed on weight (\pm 1 SEM) for the Nt (triangles), Ca (circles) and
784 NtCa (squares) lineages, when measured on *C. annuum* and *N. tabacum*.

785

786 *Figure 5.* Schematic representation of the collection of mutations obtained in the 69
787 experimentally evolved lineages. The first line represents the full TEV genome with the
788 position of the eleven mature viral proteins within the ORF. The seven other lines correspond
789 each to one evolutionary history and all the mutations obtained for this evolutionary history
790 are represented. Synonymous mutations are represented as empty circles and non-
791 synonymous mutations are represented as full circles. Mutations in black are unique whereas
792 mutations in colour are shared between several lineages. The colour corresponds to the host
793 present during the evolutionary history of all the lineages containing the shared mutation and

794 the colour code is reported on the evolutionary history names on the left. The number in
795 parentheses above each shared mutation indicates the number of times the mutation has been
796 found in the evolutionary history.

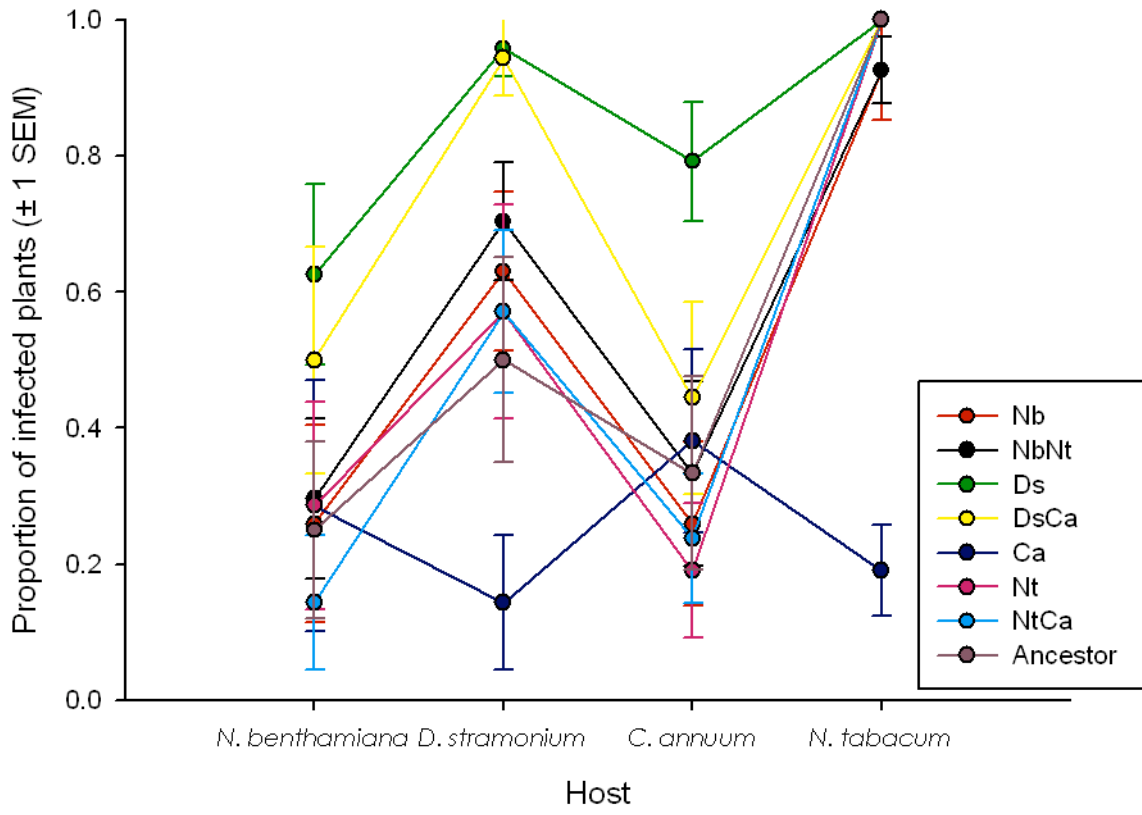
797



799

800

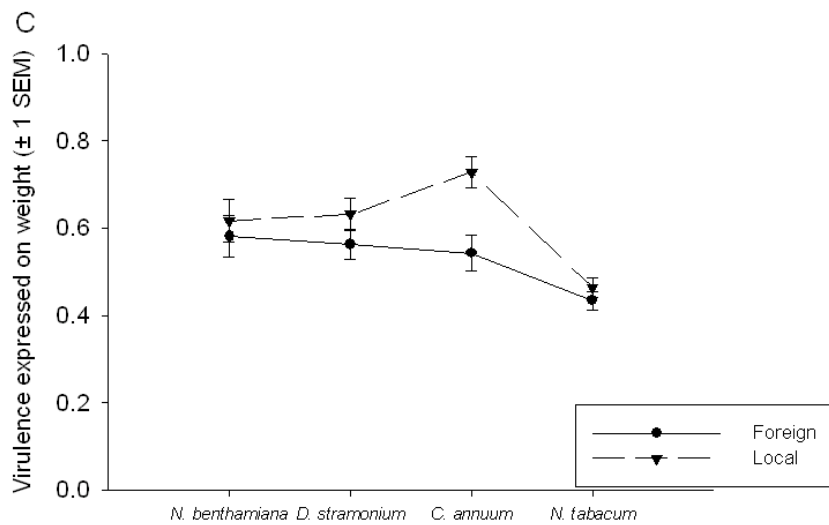
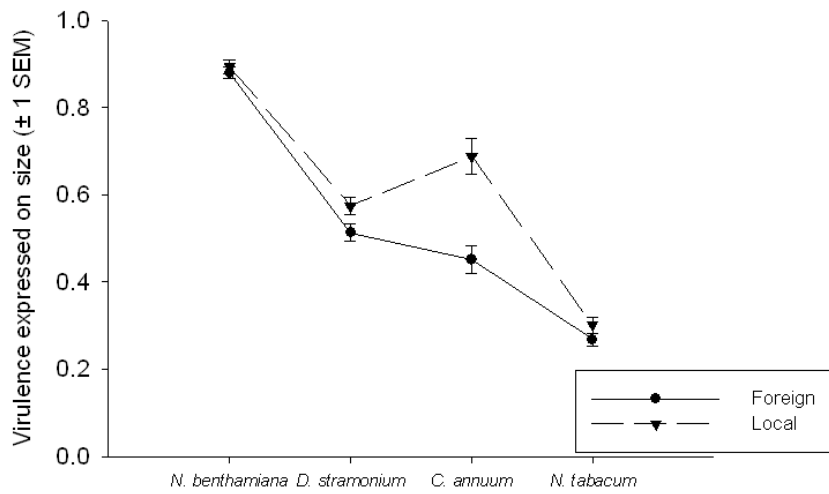
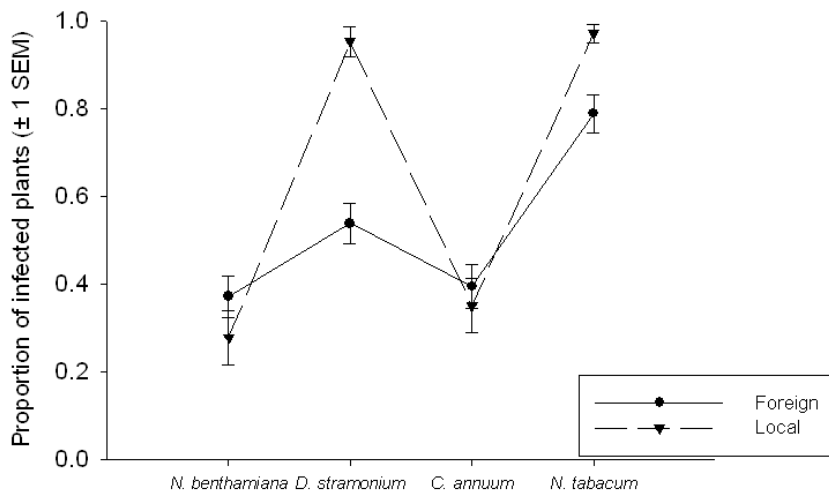
801 Figure 2



802

803

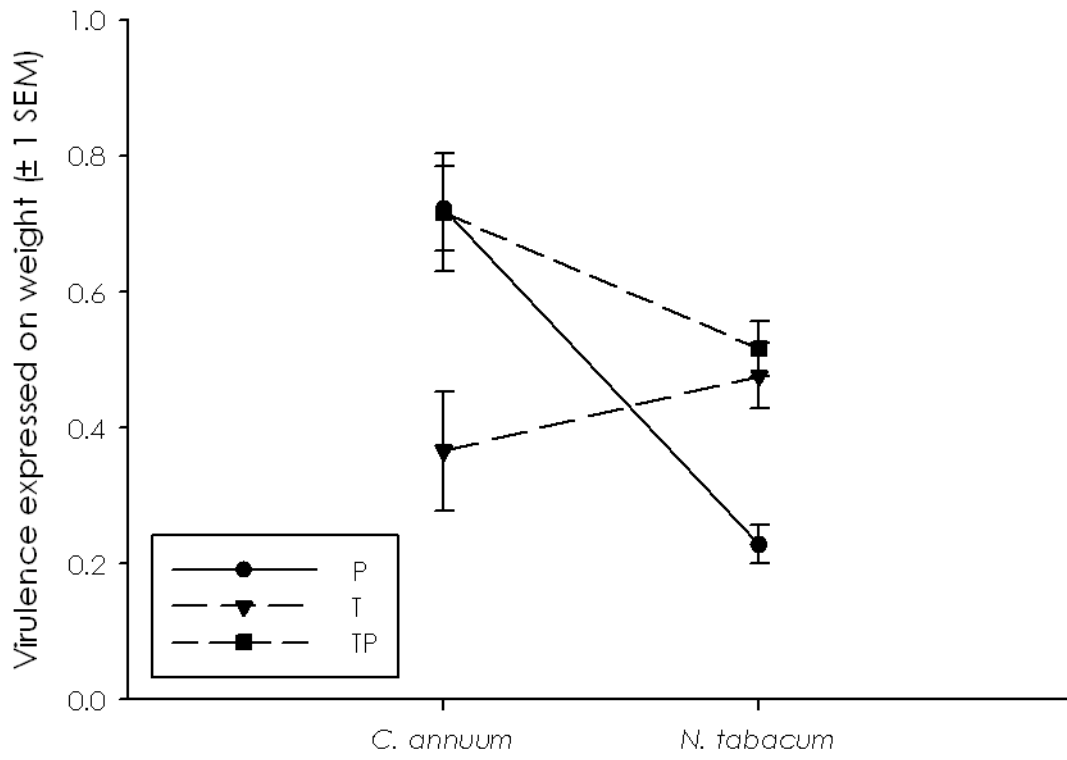
804 Figure 3



805

806

807 Figure 4



808

809

