



## Research

**Cite this article:** Lalić J, Elena SF. 2012

Epistasis between mutations is host-dependent for an RNA virus. *Biol Lett* 9: 20120396.

<http://dx.doi.org/10.1098/rsbl.2012.0396>

Received: 25 April 2012

Accepted: 21 June 2012

### Subject Areas:

evolution

### Keywords:

emerging viruses, deleterious mutations, epistasis, genotype-by-environment, virus evolution

### Author for correspondence:

Santiago F. Elena

e-mail: [santiago.elena@csic.es](mailto:santiago.elena@csic.es)

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsbl.2012.0396> or via <http://rsbl.royalsocietypublishing.org>.

One contribution to a Special Feature on 'Experimental evolution' organized by Paul Sniegowski, Thomas Bataillon and Paul Joyce.

## Evolutionary biology

# Epistasis between mutations is host-dependent for an RNA virus

Jasna Lalić<sup>1</sup> and Santiago F. Elena<sup>1,2</sup>

<sup>1</sup>Instituto de Biología Molecular y Celular de Plantas, CSIC-UPV, 46022 València, Spain

<sup>2</sup>Santa Fe Institute, Santa Fe NM 87501, USA

How, and to what extent, does the environment influence the way mutations interact? Do environmental changes affect both the sign and the magnitude of epistasis? Are there any correlations between environments in the variability, sign or magnitude of epistasis? Very few studies have tackled these questions. Here, we addressed them in the context of viral emergence. Most emerging viruses are RNA viruses with small genomes, overlapping reading frames and multifunctional proteins for which epistasis is abundant. Understanding the effect of host species in the sign and magnitude of epistasis will provide insights into the evolutionary ecology of infectious diseases and the predictability of viral emergence.

## 1. Introduction

The large majority of emerging viruses are RNA viruses [1]. However, their compact genomes comprising overlapping reading frames and multifunctional proteins and their high mutation rates may impose severe adaptive constraints [2]. Understanding the mechanistic basis of these constraints is central to explaining why some RNA viruses are more able than others to cross species boundaries. Epistasis is thought to be important in the evolution of host range [3,4]. Moreover, it has been suggested that the sign of epistasis depends on environmental severity, switching from positive to negative as environments become stressful [5]. Yet, few studies have empirically examined this possibility.

To evaluate the effect that different hosts exert on the distribution of epistatic interactions, we tested the fitness of *Tobacco etch virus* (TEV) genotypes carrying two single-nucleotide substitutions, whose independent effects were previously evaluated [6], across susceptible hosts of increasing genetic divergence from the primary host. TEV naturally infects *Solanaceae* plants, and the strain used here was isolated from *Nicotiana tabacum* [7]. Previously, we have shown that the deleterious effects of mutations were stronger as the host (i.e. the virus's environment,  $E$ ) was more genetically diverged from tobacco, and the proportion of lethal, deleterious, neutral and beneficial mutations was also altered [6]. We also found that this host dependence (i.e. plasticity or  $G \times E$ ) had two origins: antagonistic pleiotropy and changes in genetic variance for fitness across hosts [6]. Furthermore, we recently found that the fitness effect of a given mutation depended on the genetic background where it was evaluated (i.e. epistasis or  $G \times G$ ) [8]. Variation was observed both in the sign and the strength of epistasis, being negative on average and with abundant cases of reciprocal sign epistasis [8]. If  $G \times E$  and  $G \times G$  play major roles in determining TEV fitness, it is logical to expect that epistasis may also vary depending on environmental severity [9], that is, a  $G \times G \times E$  component may exist. Quantifying the extent to which  $G \times G \times E$  determines viral fitness is central to predicting the fate of viral genotypes across hosts and, ultimately, the likelihood that viruses will cross host species barriers. Epistatic interactions allowing RNA viruses to infect new hosts have been widely observed. For example, interactions between five amino acids in the coat protein of *Pelargonium flower break virus* are necessary for improving

**Table 1.** Epistasis of double mutants in each host. Average epistasis was computed after excluding lethal combinations. Sign epistasis refers to cases in which the sign of the fitness effect depends on the genetic background. Reciprocal (recip.) sign epistasis means that the sign of the fitness effect of a mutation is conditional upon the state of another locus and *vice versa*. Last row shows the significance test for the average epistasis. Red numbers indicate significant changes in epistasis from the primary host (*N. tabacum*) to alternative ones (paired *t*-tests corrected for multiple comparisons; figure 1). Errors represent  $\pm 1$  s.e.m.

genotype	<i>N. tabacum</i>	<i>D. stramonium</i>	<i>H. annuus</i>	<i>S. oleracea</i>	average epistasis
PC6/PC63	0.0730	1.5520 <sup>a</sup>	−0.0725	−0.0828	0.3674 $\pm$ 0.3965
PC6/PC76	−1.8050 <sup>a</sup> (sign)	−0.6233 (sign)	−0.1178	−0.0055	−0.6379 $\pm$ 0.4116
PC19/PC41	0.1117 <sup>a</sup> (recip. sign)	0	−0.0245	−0.0263	0.0152 $\pm$ 0.0327
PC22/PC69	−0.0293	−1.7129 <sup>a</sup>	−0.2147	−0.2106	−0.5419 $\pm$ 0.3927
PC22/PC72	0.0179	−0.3213 <sup>a</sup>	−0.2172	−0.1414	−0.1633 $\pm$ 0.0698
PC22/PC95	−1.7024 <sup>a</sup>	−0.4537	−0.1855	−0.1474	−0.6222 $\pm$ 0.3665
PC40/PC83	0.1111	−0.2108	−0.0829	−0.0535	−0.0590 $\pm$ 0.0662
PC67/PC76	0.0408	−0.5341 <sup>a</sup> (sign)	−1.0253 <sup>a</sup> (recip. sign)	0.1158	−0.3507 $\pm$ 0.2677
PC69/PC76	−1.7620 <sup>a</sup>	−0.5057 <sup>a</sup> (sign)	−0.1112	0.0221	−0.5892 $\pm$ 0.4067
PC76/PC95	0.0381	−0.5955 <sup>a</sup>	0.0127	0.0496	−0.1238 $\pm$ 0.1574
average epistasis	0.0519 $\pm$ 0.0193	−0.2834 $\pm$ 0.3187	−0.2185 $\pm$ 0.1043	−0.0480 $\pm$ 0.0316	
<i>t</i> -test (9 d.f.)	0.0358	0.4034	0.0695	0.1630	

<sup>a</sup>Epistasis significantly departs from zero within the host.

fitness in the new host *Chenopodium quinoa* [10]. Similarly, the ability of *Potato virus Y* to infect resistant pepper plants depends both on the alleles at the VPg and at the CI genes [11].

## 2. Material and methods

### (a) Virus genotypes

The 10 double mutants ([8]; electronic supplementary material, table S2) were generated by randomly combining pairs of 12 single mutations ([12]; electronic supplementary material, table S1). The particular 10 double mutants generated were randomly chosen. Mutant genotypes were generated by site-directed mutagenesis of plasmid pMTEV [7] using QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene). Infectious RNAs were obtained by *in vitro* transcription of the corresponding plasmids [13].

### (b) Host species and inoculation experiments

*Nicotiana tabacum* and *Datura stramonium* are *Solanaceae*. *Helianthus annuus* is an *Asteraceae*. *Solanaceae* and *Asteraceae* are Asterids [14]. *Spinacea oleracea* is an *Amaranthaceae*. All families are eudicots [14].

All plants were inoculated in a single block and at similar developmental stages. Nine plants per host per virus genotype were rub-inoculated at the first true leaf with 5  $\mu$ g of RNA of each genotype and 10 per cent carborundum. *Solanaceae* hosts show symptoms when infected; non-*Solanaceae* hosts do not, and infections were confirmed by RT-PCR [15]. Ten days post-inoculation (dpi), the whole infected plant, except the inoculated leaf, was collected. Tissue was frozen in liquid nitrogen and ground.

### (c) RNA purification, virus fitness and epistasis estimation

Viral RNA was purified as described elsewhere [6]. Total plant RNA concentration was measured spectrophotometrically and the samples were diluted to a final concentration of 50 ng  $\mu$ l<sup>−1</sup>.

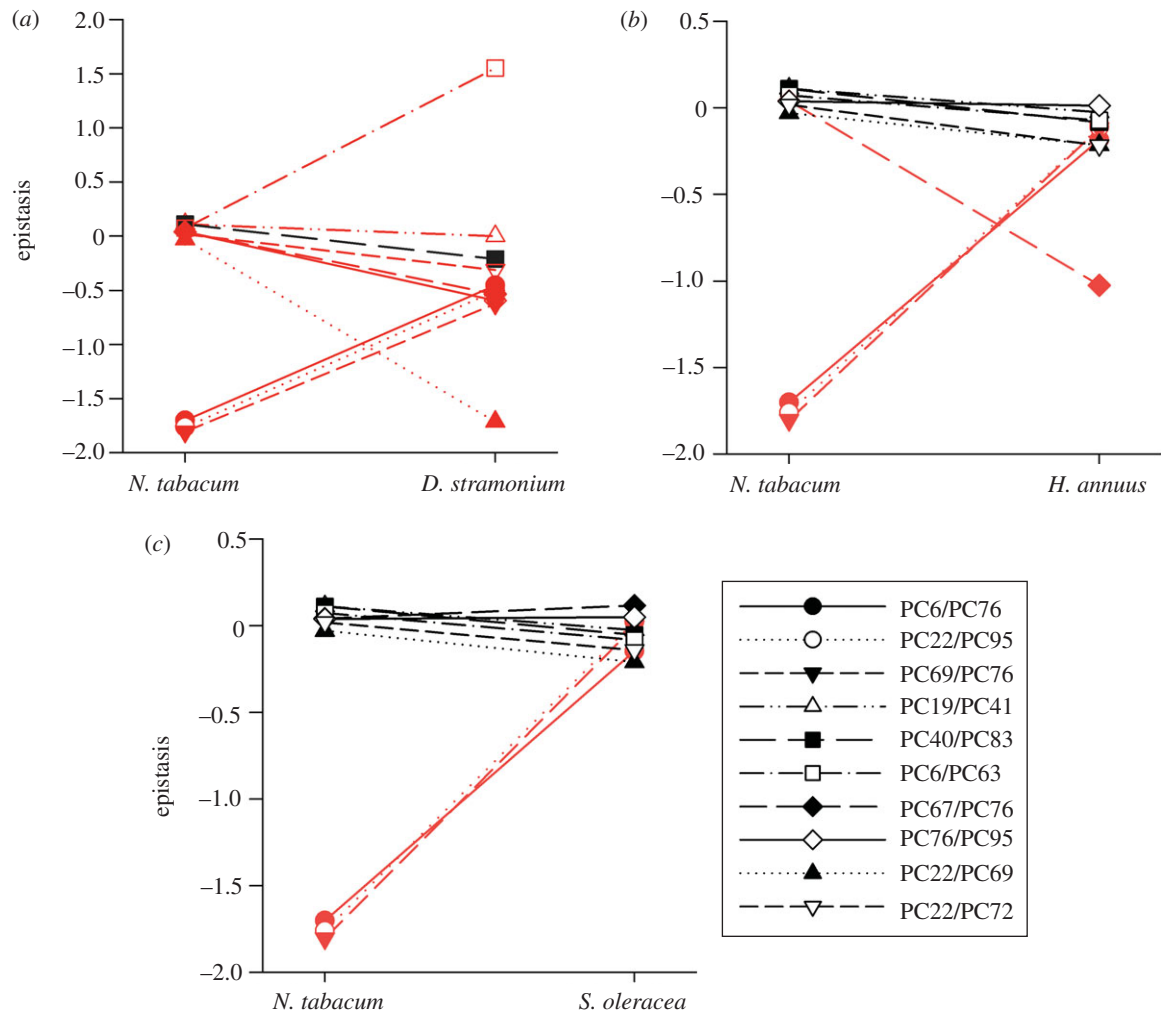
Within-plant virus accumulation was measured by absolute RT-qPCR [6].

For each genotype, a Malthusian growth rate per day was computed as  $m = 1/t \log Q_t$ , where  $Q_t$  is the number of pg of TEV RNA per 100 ng of total plant RNA quantified at  $t = 10$  dpi. Absolute fitness was defined as  $W = e^m$  (electronic supplementary material, table S1).

Epistasis between mutations  $x$  and  $y$  was calculated as  $\varepsilon_{xy} = W_{00}W_{xy} - W_{x0}W_{0y}$ , where  $W_{00}$ ,  $W_{xy}$ ,  $W_{x0}$ ,  $W_{0y}$  stand for the fitness of wild-type, double and single mutants, respectively (electronic supplementary material, table S2). Qualitatively identical results are obtained using the scaled epistasis [16].

## 3. Results

First, we sought to determine whether the number of epistatic pairs was affected by the host species. Table 1 shows the pairs of mutations evaluated on each host classified as: (i) independent effects  $\varepsilon_{xy} = 0$ , (ii) positive epistasis, and (iii) negative interactions (for each host, one-sample *t*-tests controlling for multiple comparisons). The distribution of counts for these three categories differs among hosts ( $\chi^2 = 14.157$ , 6 d.f.,  $p = 0.028$ ), with the difference being driven by an excess of non-epistatic cases in the non-*Solanaceae* (table 1). The difference is further enhanced if counts are pooled together for *Solanaceae* and non-*Solanaceae* (Fisher's exact tests,  $p = 0.003$ ). However, this classification into multiplicative versus epistatic pairs has to be taken with caution since a weak yet significant negative correlation exists between the absolute value of  $\varepsilon_{xy}$  and its error (see electronic supplementary material, table S2; partial correlation coefficient controlling for host:  $r = -0.282$ , 37 d.f., 1-tailed  $p = 0.041$ ), suggesting that the smaller the  $\varepsilon_{xy}$ , the larger its uncertainty, resulting in less power to reject the null hypothesis of independent effects.



**Figure 1.** Changes in epistasis from the primary host to alternative hosts ((a) *D. stramonium*, (b) *H. annuus* and (c) *S. oleracea*). Significant differences are indicated in red (paired *t*-tests corrected for multiple comparisons).

The above classification is just one of several possible. An alternative classification distinguishes between magnitude and sign epistasis. For magnitude epistasis, the fitness value associated with a mutation, but not its sign, changes upon the genetic background [17]. For sign epistasis, the sign of the fitness effect itself is under epistatic control [17]. Table 1 indicates which pairs match these categories. For pairs involved in significant sign epistasis, those of reciprocal type (i.e. the sign of the fitness effects change for both mutations) are also indicated. A significant difference among hosts holds if mutations are sorted according to this classification ( $\chi^2 = 14.927$ , 6 d.f.,  $p = 0.021$ ; *Solanaceae* versus non-*Solanaceae*: Fisher's exact test,  $p = 0.004$ ). With this classification scheme, the excess of independent fitness effects for non-*Solanaceae* also drives the difference among hosts. From these analyses, we can conclude that the host species has an effect on the number of epistatic interactions in TEV, with the number of independent fitness effects being significantly larger in hosts distantly related to the primary host.

Next, we identified the effect of hosts on epistasis for each pair of mutations. Figure 1 shows the change in  $\varepsilon_{xy}$  from *N. tabacum* to alternative hosts. A horizontal line means that epistasis among a pair of mutations is host-independent. Lines with positive or negative slopes indicate host-dependent epistasis. In *D. stramonium* (figure 1a), epistasis became more negative in one case, less negative in three, more positive in one and less positive in four instances. In

*H. annuus* (figure 1b), one case was significantly more negative than in tobacco and the less negative cases were the same as in *D. stramonium*. Finally, for *S. oleracea*, significant changes were detected only for the same three pervasive genotypes (table 1). Interestingly, pairs PC6/PC76, PC22/PC95 and PC69/PC76, each of which carries viable mutations when tested individually in *N. tabacum*, are not viable in this host when combined. This synthetic lethality (SL) is an extreme case of negative epistasis. However, these three genotypes are viable in the alternative hosts. By contrast, genotypes PC22/PC69 and PC67/PC76 represent cases of SL only in *D. stramonium* and *H. annuus*, respectively. These observations indicate that SL is also host-dependent. In all these cases, mutations affect different proteins (see electronic supplementary material, table S1). PC19, affecting HC-Pro, was previously described as lethal in *D. stramonium* [6], and the same lethal phenotype was observed for PC19/PC41. Conversely, PC63, affecting 6K2, also previously described as lethal in this host [6], is compensated by PC6 in protein P1, rendering a viable PC6/PC63.

When SLs are included, no host departed from the expectation of independent effects (table 1, one-sample *t*-tests;  $p \geq 0.052$ ), although significant differences among hosts exist ( $F_{3,177} = 33.660$ ,  $p < 0.001$ ). Since SLs are irrelevant in terms of evolutionary dynamics, we re-evaluated average epistasis after removing them. In this case, the average  $\varepsilon_{xy}$  becomes significantly positive in *N. tabacum* ( $p = 0.036$ ) but remains

non-significant in the alternative hosts ( $p \geq 0.070$ ). Therefore, we conclude that the intensity of epistasis decreases as the genetic divergence between the primary host and alternative hosts increases. However, this trend may be a spurious consequence of our reduced statistical power to detect small epistasis values.

The last column of table 1 shows epistasis for each double mutant averaged across hosts. A significant overall genotype effect exists ( $F_{9,177} = 168.593$ ,  $p < 0.001$ ), with epistasis ranging from negative to weakly positive. A significant genotype-by-host effect has been detected ( $F_{27,177} = 1.55 \times 10^5$ ,  $p < 0.001$ ), providing support for the importance of  $G \times G \times E$  in the architecture of viral fitness.

This ANOVA treated epistasis values as independent observations. However, this raises two statistical concerns: (i) the same mutations are involved in multiple pairs and (ii) the fitness of the wild-type on a host ( $W_{00}$ ) has been used to compute  $\varepsilon_{xy}$  for each genotype in this host. We circumvented these problems as follows: (i) the effect of using the same mutation on different combinations was removed by running the analyses for each genotype independently and making inferences valid only for each individual genotype and (ii) the non-independence introduced by re-using  $W_{00}$  was minimized using a bootstrap approach. The results from these extra analyses (see electronic supplementary material, table S3) confirm the significant  $G \times G \times E$ .

## 4. Discussion

Our experiments show that the fitness value of a given mutation depends on the genotypic background wherein it appears and on the infected host. This observation has implications for predicting the fate of viral genotypes under different and variable environments and, consequently, for the development of successful

antiviral strategies based on the use of attenuated vaccines. We stress the importance of evaluating candidate attenuating mutations in multiple genetic backgrounds and across the widest possible panel of hosts, especially in close relatives to the ones for which the vaccine is intended. Otherwise, attenuating mutations may be easily compensated by second-site changes that are viable, or even beneficial, in alternative hosts.

Our results indicate that host effects on epistasis are modulated by the degree of genetic divergence between the primary and alternative hosts. It was previously shown that point mutations had more deleterious effects as the genetic divergence from the primary host increased [6]. This observation agreed with the results of a simulation study of phage T7 showing that mutations were more severe in poor environments and milder in rich ones [5]. Furthermore, mild mutations showed negative epistasis in poor environments but weak positive epistasis in rich ones, while severe mutations showed either no epistasis or weak positive epistasis in poor environments and positive epistasis in rich ones [5]. We have shown here that epistasis was positive in the primary host (after removing SLs) but switched to no epistasis in other hosts. Together, these observations suggest that *N. tabacum* (and to a minor extent *D. stramonium*) represent rich environments for TEV, while the alternative hosts represent more stressful environments. This makes sense, considering that TEV has a coevolutionary history with *Solanaceae* hosts and thus its interaction with cellular resources and defenses is optimal. By contrast, alternative hosts may not provide the necessary resources at the right time, amount or location.

We thank Francisca de la Iglesia and Àngels Pròsper for technical assistance, Mark Zwart for help with R, Stéphanie Bedhomme and Susanna Remold for fruitful discussion. This research was supported by Ministerio de Ciencia e Innovación grant BFU2009-06993 (S.F.E.) and by CSIC JAE-Pre program (J.L.).

## References

- Holmes EC. 2009 The evolutionary genetics of emerging viruses. *Annu. Rev. Ecol. Evol. Syst.* **40**, 353–372. (doi:10.1146/annurev.ecolsys.110308.120248)
- Holmes EC. 2003 Error threshold and the constraints to RNA virus evolution. *Trends Microbiol.* **11**, 543–546. (doi:10.1016/j.tim.2003.10.006)
- Holmes EC, Rambaut A. 2004 Viral evolution and the emergence of SARS coronavirus. *Phil. Trans. R. Soc. Lond. B* **359**, 1059–1065. (doi:10.1098/rstb.2004.1478)
- Pepin KM, Wichman HA. 2007 Variable epistatic effects between mutations at host recognition sites in  $\phi$ X174 bacteriophage. *Evolution* **67**, 1710–1724. (doi:10.1111/j.1558-5646.2007.00143.x)
- You L, Yin J. 2002 Dependence of epistasis on environment and mutation severity as revealed by *in silico* mutagenesis of phage T7. *Genetics* **160**, 1273–1281.
- Lalić J, Cuevas JM, Elena SF. 2011 Effect of host species on the distribution of mutational effects for an RNA virus. *PLoS Genet.* **7**, e1002378. (doi:10.1371/journal.pgen.1002378)
- Bedoya LC, Daròs JA. 2010 Stability of *Tobacco etch virus* infectious clones in plasmid vectors. *Virus Res.* **149**, 234–240. (doi:10.1016/j.virusres.2010.02.004)
- Lalić J, Elena SF. In press. Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity*. (doi:10.1038/hdy.2012.15)
- Agrawal AF, Withlock MC. 2010 Environmental duress and epistasis: how does stress affect the strength of selection on new mutations? *Trends Ecol. Evol.* **25**, 450–458. (doi:10.1016/j.tree.2010.05.003)
- Rico P, Ivars P, Elena SF, Hernández C. 2006 Insights into the selective pressures restricting *Pelargonium flower break virus* genome variability: evidence for host adaptation. *J. Virol.* **80**, 8124–8132. (doi:10.1128/JVI.00603-06)
- Montarry J, Doumayrou J, Simon V, Moury B. 2011 Genetic background matters: a plant–virus gene-for-gene interaction is strongly influenced by genetic contexts. *Mol. Plant Pathol.* **12**, 911–920. (doi:10.1111/j.1364-3703.2011.00724.x)
- Carrasco P, de la Iglesia F, Elena SF. 2007 Distribution of fitness and virulence effects caused by single-nucleotide substitutions in *Tobacco etch virus*. *J. Virol.* **81**, 12 979–12 984. (doi:10.1128/JVI.00524-07)
- Carrasco P, Daròs JA, Agudelo-Romero P, Elena SF. 2007 A real-time RT-PCR assay for quantifying the fitness of *Tobacco etch virus* in competition experiments. *J. Virol. Meth.* **139**, 181–188. (doi:10.1016/j.jviromet.2006.09.020)
- Soltis DE, Soltis PS. 2000 Contributions of plant molecular systematics to studies of molecular evolution. *Plant Mol. Biol.* **42**, 45–75. (doi:10.1023/A:1006371803911)
- Lalić J, Agudelo-Romero P, Carrasco P, Elena SF. 2010 Adaptation of *Tobacco etch potyvirus* to a susceptible ecotype of *Arabidopsis thaliana* capacitates it for systemic infection of resistant ecotypes. *Phil. Trans. R. Soc. B* **65**, 1997–2008. (doi:10.1098/rstb.2010.0044)
- Segrè D, DeLuna A, Church GM, Kishony R. 2005 Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83. (doi:10.1038/ng1489)
- Poelwijk FJ, Tanase-Nicola S, Kiviet DJ, Tans SJ. 2011 Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol.* **272**, 141–144. (doi:10.1016/j.jtbi.2010.12.015)