biology letters
Molecular evolution

Biol. Lett. (2009) **5**, 394–396
doi:10.1098/rsbl.2008.0762
Published online 25 February 2009

Upper-limit mutation rate estimation for a plant RNA virus

Rafael Sanjuán[1,†,‡], Patricia Agudelo-Romero[1,‡] and Santiago F. Elena[1,2,*]

[1]*Instituto de Biología Molecular y Celular de Plantas, CSIC-UPV, Campus UPV CPI 8E, Ingeniero Fausto Elio s/n, 46022 València, Spain*
[2]*The Santa Fe Institute, Santa Fe, NM 87501, USA*
**Author for correspondence (sfelena@ibmcp.upv.es).*
[†]*Present address: Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, 46980 València, Spain.*
[‡]*These authors contributed equally to this work.*

It is generally accepted that mutation rates of RNA viruses are inherently high due to the lack of proofreading mechanisms. However, direct estimates of mutation rate are surprisingly scarce, in particular for plant viruses. Here, based on the analysis of *in vivo* mutation frequencies in *tobacco etch virus*, we calculate an upper-bound mutation rate estimation of 3×10^{-5} per site and per round of replication; a value which turns out to be undistinguishable from the methodological error. Nonetheless, the value is barely on the lower side of the range accepted for RNA viruses, although in good agreement with the only direct estimate obtained for other plant viruses. These observations suggest that, perhaps, differences in the selective pressures operating during plant virus evolution may have driven their mutation rates towards values lower than those characteristic of other RNA viruses infecting bacteria or animals.

Keywords: experimental evolution; mutation rate; plant virus; virus evolution

1. INTRODUCTION

The rate of spontaneous mutation is a key evolutionary parameter, yet accurate estimations are available only for a handful of organisms. One of the most widely accepted differences in mutation rates is the one between RNA viruses and dsDNA organisms, the former mutating orders of magnitude more frequently than the latter. Such elevated mutation rates, though not necessarily evolved in response to natural selection (Elena & Sanjuán 2005; Clune *et al.* 2008), can explain why RNA virus show rapid evolution at the molecular level, although other factors such as replication speed, transmission mode or genomic architecture have to be considered as well (Duffy *et al.* 2008). However, our current knowledge of RNA virus mutability is mainly based on work with animal viruses and bacteriophages (Elena *et al.* 2008).

High genetic stability has frequently been reported for plant RNA viruses (Rodríguez-Cerezo *et al.* 1991; Fraile *et al.* 1997; Marco & Aranda 2005; Herránz *et al.* 2008), although substitution rates within the range of animal RNA viruses have also been reported (Fargette *et al.* 2008; Gibbs *et al.* 2008). This peculiar behaviour might be due to stronger selective constraints acting on plant viruses, weaker immune-mediated positive selection (García-Arenal *et al.* 2001), the existence of strong bottlenecks during cell-to-cell movement and systemic colonization of distal tissues (Hall *et al.* 2001*a*; Sacristán *et al.* 2003; Li & Roossink 2004), more severe bottlenecks during horizontal aphid-mediated transmission (Ali *et al.* 2006; Betancourt *et al.* 2008), or differences in replication mode compared with lytic viruses (French & Stenger 2003). Another, more obvious, possibility is that plant viruses show reduced rates of spontaneous mutation. However, the only available direct mutation rate estimation for a plant virus is for the *tobacco mosaic virus* (TMV; Malpica *et al.* 2002). In this pioneer study, mutations at the viral movement protein (MP) gene were sampled from TMV populations infecting transgenic *Nicotiana tabacum* plants that expressed the MP protein, hence minimizing selection against TMV MP mutants. A mutation rate of 2.4×10^{-5} mutations/base/replication round (m/b/r) was obtained, which lies on the lower side of the range usually accepted for animal riboviruses (Drake & Holland 1999), although this estimate was subject to uncertainties related to the number of infection cycles elapsed during the mutation-accumulation phase and the fraction of mutations that produced a selectable phenotype.

To further investigate whether plant viruses show unusually low rates of mutation, we followed a strategy that allowed us to obtain an upper-bound estimation for the mutation rate, based on the quantification of genetic variability in *tobacco etch virus* (TEV), a positive-stranded RNA virus belonging to the Potyviridae family. Then, we compared this estimate with those available in the literature.

2. MATERIAL AND METHODS

(a) *Viral populations*

Chenopodium quinoa leaves were inoculated with a TEV extract and six necrotic local lesions were isolated. This procedure (the equivalent of plaque-picking in lytic viruses) introduces a severe bottleneck in the viral population, effectively eliminating genetic variation (de la Iglesia & Elena 2007). Each of these isolates was used to infect a different four-week old *N. tabacum* plant and, following the appearance of systemic symptoms, virions were extracted from whole plants as previously described (Carrasco *et al.* 2007). Any mutation detected would necessarily be generated after the bottleneck.

(b) *Molecular methods*

The region encompassing nucleotides 7808–9437, which corresponds to the 5′ end of the viral polymerase gene *NIb*, the entire capsid protein gene and the proximal portion of the 3′ untranslated region, was sequenced. *Moloney murine leukemia virus* (MMLV) reverse transcriptase (RT) (Fermentas) was used to obtain cDNA from the viral extract and the high-fidelity Phusion DNA polymerase (Finnzymes) was used to obtain a PCR fragment, which was subsequently gel-purified, ligated to a pUC19/SmaI (Fermentas) and used for transformation of *Escherichia coli* DH5α. Plasmid DNA was purified and sequenced using the ABI prism big dye terminator cycle sequencing kit v. 3.1 (Applied Biosystems), vector-based primers and internal primers. Labelled products were resolved in an ABI 3100 genetic analyser (Applied Biosystems).

(c) *Estimating mutation rate from mutant frequencies*

Following Drake & Holland (1999), the frequency of mutations per base f after c cycles of cell infection is given by $f=2c\mu$, where μ is

One contribution of 11 to a Special Feature on 'Whole organism perspectives on understanding molecular evolution'.

Received 11 December 2008
Accepted 16 January 2009

Table 1. Upper-limit estimate for the mutation rate (μ_{max}) for several RNA plant viruses.

virus[a]	host	μ_{max} ($\pm$s.e.m.)$\times 10^{-4}$	Reference
CMV	*Capsicum annuum*	15.34 ± 0.71	Schneider & Roossinck (2001)
	C. annuum	1.39 ± 0.07	Pita *et al.* (2007)
	Nicotiana benthamiana	6.64 ± 0.95	Schneider & Roossinck (2000)
	Nicotiana tabacum	0.20 ± 0.09	Pita *et al.* (2007)
CCMV	*N. benthamiana*	5.29 ± 4.93	Schneider & Roossinck (2000)
TMV	*C. annuum*	11.02 ± 0.12	Schneider & Roossinck (2001)
	Collinsia heterophylla	4.74	Kearney *et al.* (1999)
	Fagopyrus esculentum	4.55	Kearney *et al.* (1999)
	Lycopersicum esculentum	1.45 ± 0.51	Schneider & Roossinck (2001)
	N. benthamiana	4.21 ± 0.69	Schneider & Roossinck (2000)
	N. tabacum	4.14	Kearney *et al.* (1999)
	N. tabacum	0.24 ± 0.00	Malpica *et al.* (2002)[b]
	Phacelia campanularia	16.81	Kearney *et al.* (1999)
	Plantago sp.	8.50	Kearney *et al.* (1999)
	Solanum nigrum	4.21	Kearney *et al.* (1999)
	Tagetes erecta	8.15	Kearney *et al.* (1999)
WSMV	*Zea mays*	9.01 ± 0.90	Hall *et al.* (2001*b*)

[a]CMV, *cucumber mosaic virus*; CCMV, *cowpea chlorotic mottle virus*; TMV, *tobacco mosaic virus*; WSMV, *wheat streak mosaic virus*.
[b]This study reported a μ estimation instead of the μ_{max} upper limit.

the mutation rate per base and copying event and '2' reflects the fact that offspring genomic strands are produced via the synthesis of intermediary antigenomic strains. For lethal mutations, $c=1$ and thus a direct estimation of μ can be obtained. However, since c is unknown in our experiment, f can only be used to obtain an upper-limit estimator of the real mutation rate ($\mu < f = \mu_{max}$). We estimate μ_{max} in samples from TEV-infected plants by high-fidelity RT-PCR, molecular cloning and sequencing.

3. RESULTS AND DISCUSSION

Out of 83 clones and a total of 135 207 nucleotides sequenced, only four mutations were observed, yielding a mutation rate ($\pm$s.e.m.) of $\mu_{max} = (2.96 \pm 0.32) \times 10^{-5}$ m/b/r. Two mutations were found in the *NIb* gene, one synonymous (C8029U) and the other (C8248U) with a probable major effect on the protein folding, since it replaces P2750 by S, an amino acid with a small polar radical. Two other non-synonymous mutations were found in the CP gene, C8980A (Q2994K) and C9008A (A3003D). The first mutation replaced a positively charged side chain by another one with similar characteristics. The second one may have a stronger effect on fitness since it replaced a small apolar radical by a larger side chain with net negative charge. Assuming that only P2750S and A3003D may be lethal for the virus, then the estimated mutation rate would be $\mu = 1.48 \times 10^{-5}$ m/b/r. This value is in good agreement with the indirect estimate reported by Carrasco *et al.* (2007) for the deleterious genomic mutation rate.

The use of a high-fidelity DNA polymerase minimizes the probability that the observed mutations were due to PCR errors, but reverse transcription artefacts are more likely since the error rate of MMLV RT is probably higher than 3.3×10^{-5} m/b/r (Arezi & Hogrefe 2007). Therefore, we can only safely conclude that the obtained μ_{max}, at least for the sequenced region, in the particular host employed and experimental conditions, is as low as that for a common RT enzyme.

How do our data compare with the levels of genetic variability reported for other plant RNA viruses? Table 1 compiles μ_{max} for four different viruses on different hosts. These values were in all cases estimated after characterizing the mutant spectrum from individual plants. The median μ_{max} across viruses and experiments is 4.74×10^{-4} m/b/r, which is in the range of estimated rates for animal RNA viruses and some RNA bacteriophages (Drake & Holland 1999). It is worth noting that estimates are homogeneous for the five different viruses (two-way non-parametric ANOVA: $H = 1.140$, d.f. $= 2$, $p = 0.566$), although they significantly vary across host plants ($H = 22.424$, d.f. $= 9$, $p = 0.008$), suggesting that virus mutation rates depend on the environment in which they replicate, as suggested by Pita *et al.* (2007).

The approximately 16-fold lower μ_{max} obtained for TEV compared with the above median value can be attributed to several factors. It is possible that TEV shows an unusually low rate of spontaneous mutation or that tobacco plants are hosts in which mutation rate is particularly low. It must also be noted that the region we have sequenced contains *cis*-acting secondary structure motifs necessary for genome replication (Haldeman-Cahill *et al.* 1998), and mutations affecting these motifs render replication-incompetent viruses. Therefore, our μ_{max} value might be closer to the actual mutation rate than μ_{max} values obtained in studies where a less-constrained region was sequenced. Hence, most of the genetic variation sampled in the studies are listed in table 1, and in previous work with animal RNA viruses and bacteriophages might correspond to neutral or nearly neutral mutations that had been segregating in the population for long periods of time, thus reaching relatively high frequencies, or that some of these mutations were beneficial. In our experiments, by contrast, long-term variation could not be present because infections were initiated from a single lesion-forming unit. Furthermore, the virus used in our experiments had been replicating in tobacco plants for 30 passages prior to the assays without showing significant changes in several fitness traits (data not shown), suggesting that the mutations sampled were probably not beneficial.

Our data allow us to conclude that the mutation rate of TEV is in the lower range of those reported for other RNA viruses, but, obviously, we cannot conclude that plant RNA viruses show low mutation rates in general. However, the fact that our estimate is very close to the only previous estimate available for a plant RNA virus (Malpica *et al.* 2002) at least suggests so. Finally, slow rates of molecular evolution in plant viruses could be partly explained by low mutation rates, but the complex relationships between mutation and evolution need to be considered in detail to explain why some plant viruses show unusually high levels of genetic stability while others evolve fast.

We thank Francisca de la Iglesia for technical assistance. This work was supported by grant BFU2006-14819-C02-01/BMC from the Spanish MICINN. R.S. also acknowledges current support from MICINN grant BFU2008-03978/BMC and the Ramón y Cajal programme.

Ali, A., Li, H., Schneider, W. L., Sherman, D. J., Gray, S., Smith, D. & Roossinck, M. J. 2006 Analysis of genetic bottlenecks during horizontal transmission of *cucumber mosaic virus*. *J. Virol.* **80**, 8345–8350. (doi:10.1128/JVI.00568-06)

Arezi, B. & Hogrefe, H. H. 2007 *Escherichia coli* DNA polymerase III epsilon subunit increases *Moloney murine leukemia virus* reverse transcriptase fidelity and accuracy of RT-PCR procedures. *Anal. Biochem.* **360**, 84–91. (doi:10.1016/j.ab.2006.10.009)

Betancourt, M., Fereres, A., Fraile, A. & García-Arenal, F. 2008 Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus. *J. Virol.* **82**, 12 416–12 421. (doi:10.1128/JVI.01542-08)

Carrasco, P., de la Iglesia, F. & Elena, S. F. 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)

Clune, J., Misevic, D., Ofria, C., Lenski, R. E., Elena, S. F. & Sanjuán, R. 2008 Natural selection fails to optimize mutation rates in rugged fitness landscapes. *PLoS Comput. Biol.* **4**, e1000187. (doi:10.1371/journal.pcbi.1000187)

de la Iglesia, F. & Elena, S. F. 2007 Fitness declines in *tobacco etch virus* upon serial bottleneck transfers. *J. Virol.* **81**, 4941–4947. (doi:10.1128/JVI.02528-06)

Drake, J. W. & Holland, J. J. 1999 Mutation rates among RNA viruses. *Proc. Natl Acad. Sci. USA* **96**, 13 910–13 913. (doi:10.1073/pnas.96.24.13910)

Duffy, S., Shackelton, L. A. & Holmes, E. C. 2008 Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**, 267–276. (doi:10.1038/nrg2323)

Elena, S. F. & Sanjuán, R. 2005 On the adaptive value of high mutation rates in RNA viruses: separating causes from consequences. *J. Virol.* **79**, 11 555–11 558. (doi:10.1128/JVI.79.18.11555-11558.2005)

Elena, S. F., Agudelo-Romero, P., Carrasco, P., Codoñer, F. M., Martin, S., Torres-Barceló, C. & Sanjuán, R. 2008 Experimental evolution of plant RNA viruses. *Heredity* **100**, 478–483. (doi:10.1038/sj.hdy.6801088)

Fargette, D. *et al.* 2008 Rice yellow mottle virus, an RNA plant virus, evolves as rapidly as most RNA animal viruses. *J. Virol.* **82**, 3584–3589. (doi:10.1128/JVI.02506-07)

Fraile, A., Escriu, F., Aranda, M. A., Malpica, J. M., Gibbs, A. J. & García-Arenal, F. 1997 A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *J. Virol.* **71**, 8316–8320.

French, R. & Stenger, D. C. 2003 Evolution of *wheat streak mosaic virus*: dynamics of population growth within plants may explain limited variation. *Annu. Rev. Phytopathol.* **41**, 199–214. (doi:10.1146/annurev.phyto.41.052002.095559)

García-Arenal, F., Fraile, A. & Malpica, J. M. 2001 Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* **39**, 157–186. (doi:10.1146/annurev.phyto.39.1.157)

Gibbs, A. J., Ohshima, K., Phillips, M. J. & Gibbs, M. J. 2008 The prehistory of potyviruses: their initial radiation was during the dawn of agriculture. *PLoS ONE* **3**, e2523. (doi:10.1371/journal.pone.0002523)

Haldeman-Cahill, R., Daròs, J. A. & Carrington, J. C. 1998 Secondary structures in the capsid protein coding sequence and 3′ nontranslated region involved in amplification of the *tobacco etch virus* genome. *J. Virol.* **72**, 4071–4079.

Hall, J. S., French, R., Hein, G. L., Morris, J. & Stenger, D. C. 2001*a* Three distinct mechanisms facilitate genetic isolation of sympatric *wheat streak mosaic virus* lineages. *Virology* **282**, 230–236. (doi:10.1006/viro.2001.0841)

Hall, J. S., French, R., Morris, T. J. & Stenger, D. C. 2001*b* Structure and temporal dynamics of populations within *wheat streak mosaic virus* isolates. *J. Virol.* **75**, 10 231–10 243. (doi:10.1128/JVI.75.21.10231-10243.2001)

Herránz, M. C., Al Rwahnih, M., Sánchez-Navarro, J. A., Elena, S. F., Choueiri, E., Myrta, A. & Pallás, V. 2008 Low genetic variability in the coat and movement proteins of *American plum line pattern virus* isolates from different geographic origins. *Arch. Virol.* **153**, 367–373. (doi:10.1007/s00705-007-1100-4)

Kearney, C. M., Thomson, M. J. & Roland, K. E. 1999 Genome evolution of *tobacco mosaic virus* populations during long-term passaging in a diverse range of hosts. *Arch. Virol.* **144**, 1513–1526. (doi:10.1007/s00705 0050607)

Li, H. & Roossinck, M. J. 2004 Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J. Virol.* **78**, 10 582–10 587. (doi:10.1128/JVI.78.19.10582-10587.2004)

Malpica, J. M., Fraile, A., Moreno, I., Obies, C. I., Drake, J. W. & García-Arenal, F. 2002 The rate and character of spontaneous mutation in an RNA virus. *Genetics* **162**, 1505–1511.

Marco, C. F. & Aranda, M. A. 2005 Genetic diversity of a natural population of *cucurbit yellow stunting disorder virus*. *J. Gen. Virol.* **86**, 815–822. (doi:10.1099/vir.0.80584-0)

Pita, J. S., de Miranda, J. R., Schneider, W. L. & Roossinck, M. J. 2007 Environment determines fidelity for an RNA virus replicase. *J. Virol.* **81**, 9072–9077. (doi:10.1128/JVI.00587-07)

Rodríguez-Cerezo, E., Elena, S. F., Moya, A. & García-Arenal, F. 1991 High genetic stability in natural populations of the plant RNA virus *tobacco mild green mosaic virus*. *J. Mol. Evol.* **32**, 328–332. (doi:10.1007/BF02102191)

Sacristán, S., Malpica, J. M., Fraile, A. & García-Arenal, F. 2003 Estimation of population bottlenecks during systemic movement of *tobacco mosaic virus* in tobacco plants. *J. Virol.* 77, 9906–9911. (doi:10.1128/JVI.77.18.9906-9911.2003)

Schneider, W. L. & Roossinck, M. J. 2000 Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host. *J. Virol.* **74**, 3130–3134. (doi:10.1128/JVI.74.7.3130-3134.2000)

Schneider, W. L. & Roossinck, M. J. 2001 Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *J. Virol.* **75**, 6566–6571. (doi:10.1128/JVI.75.14.6566-6571.2001)