# The evolutionary genetics of emerging plant RNA viruses.

Santiago F Elena, Stéphanie Bedhomme, Purificación Carrasco, José M Cuevas, Francisca de La Iglesia, Guillaume Lafforgue, Jasna Lalić, Angels Pròsper, Nicolas Tromas, Mark P Zwart

MPMI Vol. 24, No. 3, 2011, pp. 287–293. doi:10.1094/MPMI-09-10-0214. © 2011 The American Phytopathological Society

*CURRENT REVIEW*

# The Evolutionary Genetics of Emerging Plant RNA Viruses

**Santiago F. Elena,[1,2] Stéphanie Bedhomme,[1] Purificación Carrasco,[1] José M. Cuevas,[1] Francisca de la Iglesia,[1] Guillaume Lafforgue,[1] Jasna Lalić,[1] Àngels Pròsper,[1] Nicolas Tromas,[1] and Mark P. Zwart[1]**

[1]Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-UPV, Campus UPV CPI 8E, Ingeniero Fausto Elio s/n, 46022 València, Spain; [2]The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, U.S.A.

Submitted 22 September 2010. Accepted 9 November 2010.

**Over the years, agriculture across the world has been compromised by a succession of devastating epidemics caused by new viruses that spilled over from reservoir species or by new variants of classic viruses that acquired new virulence factors or changed their epidemiological patterns. Viral emergence is usually associated with ecological change or with agronomical practices bringing together reservoirs and crop species. The complete picture is, however, much more complex, and results from an evolutionary process in which the main players are ecological factors, viruses' genetic plasticity, and host factors required for virus replication, all mixed with a good measure of stochasticity. The present review puts emergence of plant RNA viruses into the framework of evolutionary genetics, stressing that viral emergence begins with a stochastic process that involves the transmission of a preexisting viral strain into a new host species, followed by adaptation to the new host.**

A rigorous definition of an emerging virus would be "the causal agent of an infectious disease of viral aetiology whose incidence is increasing following its first introduction into a new host population or whose incidence is increasing in an existing host population as a result of long-term changes in its underlying epidemiology" (Woolhouse and Dye 2001), and we will add "often accompanied by a significant increase in symptom severity" (Cleaveland et al. 2007). Accordingly, the epidemic spread 20 years ago of necrogenic strains of *Cucumber mosaic virus* (CMV) on tomato crops in eastern Spain (Escriu et al. 2000) or the worldwide ongoing epidemic of *Pepino mosaic virus* in tomatoes should both be considered as paradigms of emerging viral infection.

Emerging viruses come from host species in which they are already established and which play the role of a reservoir host during emergence. Species jumps, or spillovers, have given rise to devastating epidemics, as exemplified above, but there are numerous examples of species jumps that have had far less dramatic consequences. There are even many viruses that have a long history of routinely jumping between species without triggering major epidemics (Thresh 2006).

The study of viral emergence can be divided into three phases. The first phase (I) accounts for the mechanisms and limitations involved in jumping the species barrier. The second phase (II) includes the subsequent evolutionary dynamics that lead to a virus well adapted to its new host. The third phase (III) comprises the epidemiological spread of this well-adapted virus in the new host population. A detailed description of these three phases is beyond the scope (and length) of this review. Therefore, we will only elaborate on the evolutionary genetic principles underlying phases I and II. Nevertheless, this division in phases is somewhat arbitrary, since, as we will see below, some of the mechanisms operate during more than one phase.

We will focus this review entirely on RNA viruses (including pararetroviruses) because of their great evolvability, a consequence of combining highly error-prone replication, large population sizes, and rapid replication rates (Elena and Sanjuán 2008). Some of our conclusions may also be valid for fast-evolving DNA viruses, such as geminiviruses.

## Genetic determinants of phase I.

The first step in viral emergence is the exposure of the new host species to the virus (Fig. 1). The rate of exposure will be a function of the ecology of the two hosts and of the transmission biology of the virus, including any relevant vectors. The initial infection of individuals of the new host species is a pivotal step in emergence. However, most viruses transferred to new hosts replicate poorly and are inefficiently transmitted. Therefore, the preexistence of host-range mutants within the standing genetic variation in the reservoir host increases the probabilities of a successful jump. The amount of standing genetic variation would depend i) on the rates of mutation and recombination, ii) on the distribution of mutational effects on viral fitness, and iii) on the strength of genetic drift and gene flow among subpopulations.

## Rates of mutation in plant RNA viruses.

Figure 2 shows estimates of mutation rates for CMV, *Cowpea chlorotic mottle virus*, *Chrysanthemum chlorotic mottle viroid* (CChMVd), *Tobacco etch virus* (TEV), *Tobacco mosaic virus* (TMV), and *Wheat streak mosaic virus* (WSMV) on different hosts. All values shown in Figure 2 for RNA viruses were estimated by evaluating the genetic variability found in plants infected with inocula containing no genetic variability and, thus, correspond to upper-limit estimates (Sanjuán et al. 2009). In the case of CChMVd, the estimate corresponds to the actual mutation rate, as estimated from the frequency of lethal alleles in the population (Gago et al. 2009). The main observation that can be drawn is the existence of homogeneity in mutation rates among plant RNA viruses, with values ranging over less than one order of magnitude. The viroid CChMVd,

Corresponding author: Santiago F. Elena; E-mail: sfelena@ibmcp.upv.es

however, shows a mutation rate that is clearly larger than the values observed for RNA viruses, as expected for an organism with a much smaller genome (Gago et al. 2009). All other (ecological and genetic) parameters being equal, viruses with high mutation rates are more likely to generate host-range mutants and emergent epidemics, although this hypothetical relationship has not yet been investigated experimentally. In addition, it is important to note that host interference with replication fidelity can influence mutation rates (Pita et al. 2007).

**Recombination in plant RNA viruses.**

Recombination potentially increases fitness by creating advantageous genotypes and removing deleterious mutations, suggesting that will bolster the process of emergence. However, this possibility is still controversial. While some studies have proclaimed that recombination may assist the process of cross-species transmission (Chare and Holmes 2006; Codoñer and Elena 2008), others have pointed out that the association between recombination and emergence is circumstantial (Holmes 2008). The vast majority of references illustrating examples of recombinant genotypes among plant viruses are based on the analyses of epidemiological sequence data (Awadalla 2003). Phylogenetic data have at least one major drawback; they do not represent an unbiased sample of all recombination events but only epitomize successful recombinant genotypes sorted out by natural selection or those genotypes that generally induce new pathologies.

The estimates reported for in vivo recombination rates were obtained for *Brome mosaic virus* (BMV) (Bruyere et al. 2000) and for the double-stranded DNA pararetrovirus *Cauliflower mosaic virus* (CaMV) (Froissart et al. 2005). Bruyere and associates (2000) engineered four BMV RNA3 containing artificial markers scattered along it. Pairs of these markers were coinoculated on *Chenopodium quinoa* leaves and the progeny were analyzed. The majority of doubly infected lesions contained recombinant RNA3. García-Arenal and associates (2001) analyzed the data corresponding to two of these pairs and estimated that the recombination rate was in the range of $1.07 \times 10^{-4}$ to $1.76 \times 10^{-4}$ per base. Later, Urbanowicz and associates (2005) extended the study to BMV RNA1 and RNA2 and found that, although the rate of recombination varied across different RNA regions, on average, the number of crossover events per segment and replication cycle was about one. Froissart and associates (2005) found that 21 days after stoichiometric inocu-

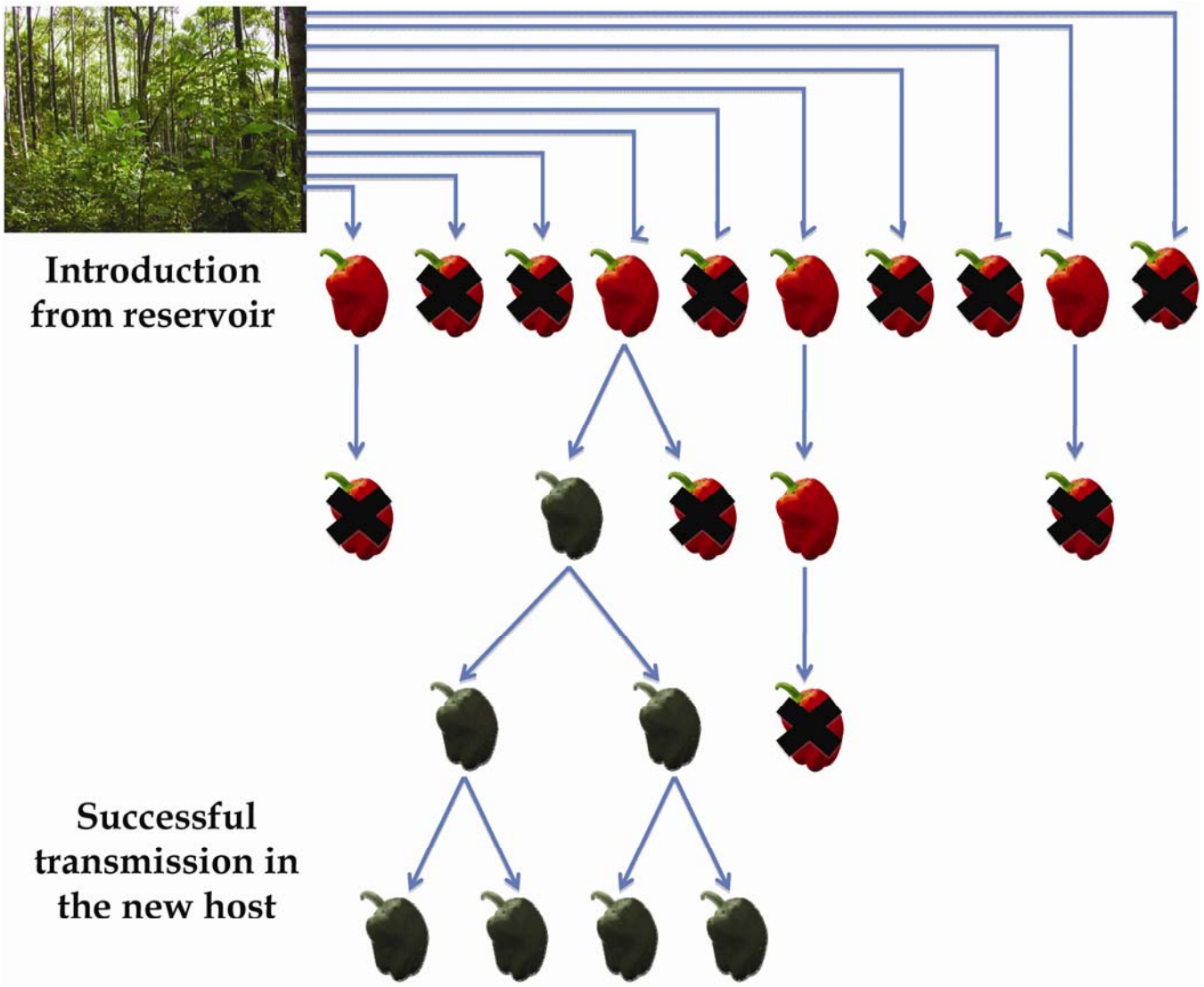


**Fig. 1.** Host-switching process. In a first phase, the virus jumps from its natural host to the new one. Each arrow departing from the forest and ending in the pepper crop represents an independent spillover. Most of these transmissions will not produce a successful infection (black crosses). In a few cases, the virus will replicate sufficiently enough to be transmitted for a second time or even a third time, but without triggering an epidemic. It is a very rare event when a virus increases its fitness in the new host, allowing for its successful transmission and, therefore, becoming epidemic (here indicated by the change in color).

lation with a wild-type CaMV and viruses carrying neutral mutations, half of the sequenced CaMV genotypes were recombinant. The calculated recombination rate falls in the range $2 \times 10^{-5}$ to $4 \times 10^{-5}$ per base and replication cycle. All together, these estimates are roughly in the same order of magnitude as the estimates of mutation rates shown in Figure 2. This coincidence suggests that recombination is a source of variation, perhaps as important as mutation.

Recombination rates are controlled by two factors, the ability of the viral replicases to undergo template switching and the multiplicity of infection (MOI) during infection. The first factor clearly varies among viruses as a function of their biology and, for example, negative-strand RNA viruses are expected to be less recombinogenic because their RNA is never naked (Chare et al. 2003). The second factor depends on the peculiarities of each virus-host pair and has started receiving attention only very recently. González-Jara and associates (2009) evaluated the frequency of TMV multiple infections within a single infected *Nicotiana benthamiana*. These authors tracked the kinetics of infection of two different TMV genotypes (respectively labeled with two fluorescent proteins) by counting the number of singly and coinfected cells. Their results indicate that MOI can be as high as 6 during initial stages of infection, although the value decreased as infection progressed down to 1 or 2, both in inoculated and systemic leaves. Gutiérrez and associates (2010) have also explored the evolution of MOI during CaMV infection of turnips. These authors found that MOI is a hump-shaped function of time; it was low (approximately 2) early after infection but increased to a maximum of approximately 12 at intermediate times and then declined again. Miyashita and Kishino (2010) have also performed coinoculation experiments with *Soil-borne wheat mosaic virus* carrying RNA2 pairs labeled with two different fluorescent proteins. These authors focused their study on evaluating the MOI during the initial stages of cell-to-cell movement from the initial infectious foci. In these experiments, MOI was approximately 6 for the movement from the initially infected cells and significantly reduced to 5 after the second movement. Additionally, Miyashita and Kishino (2010) observed that, within coinfected foci, most cells were coinfected, but some only showed one type of fluorescence. As time went on and the foci expanded, this initial aggregation turned into spatial separation of the two colors. This result is consistent with the observation that, during mixed infections, different genotypes of the same potyvirus exclude each other (Dietrich and Maiss 2003). This spatial structuring of genotypes during the course of infection, together with the relatively low MOI values (in the order of units), could effectively diminish the possibilities for different genotypes to coinfect cells and recombine.

**Fitness tradeoffs across hosts.**

A fundamental challenge for host-switching viruses is that different hosts impose different selective requirements for viruses; so acquiring the ability to replicate in a new one may impose a fitness burden in the original. These fitness tradeoffs can be generated by different mechanisms, antagonistic pleiotropy (AP) being the simplest and most intuitive one. AP means that mutations that are beneficial in one host may be deleterious in an alternative one. A second mechanism that promotes tradeoffs results from mutation accumulation by genetic drift. Accumulated mutations may be neutral in the current host but may be essential in a future one (Kawecki 1994). Although both mechanisms involve differences in mutational effects across hosts, it is necessary to stress that they are by no means equivalent phenomena. While natural selection is the only reason for the tradeoff in the former, genetic drift is important in the latter. Most of the accumulated evidence suggests that AP is the principal but not the only reason for fitness tradeoffs (Elena et al. 2009). AP may be an unavoidable consequence of the small size of viral genomes, which in many instances contain overlapping genes and encode multifunctional proteins, making it extremely difficult to optimize one function without jeopardizing another. Fitness tradeoffs across alternative hosts have been reported for several plant viruses. For instance, Jenner and associates (2002) found that *Turnip mosaic virus* (TuMV) capable of infecting two different genotypes of turnips paid a fitness penalty compared with the ancestral virus, which was only capable of infecting a given genotype. Similarly, Wallis and associates (2007) have shown that, following serial passages in peas, *Plum pox virus* increased infectivity, viral load, and virulence in the new host with a concomitant reduction in transmission efficiency in the original host peach trees. As a last example, Agudelo-Romero and associates (2008c) found that TEV lineages adapted to pepper showed reduced virulence in tobacco, while serial passages in tobacco did not affected the virulence in pepper.

Some pieces of evidence also suggest that the fitness of a virus simultaneously facing multiple hosts is either constrained by the most restrictive one or is not subject to a tradeoff at all. In this respect, theory predicts that the extent to which multihost viruses evolve depends on the frequency at which viruses transmit among heterologous hosts (Wilke et al. 2006). When transmission among heterologous hosts represents an infrequent event, the viral population essentially adapts to the current one. However, if heterologous transmissions are frequent, the viral population behaves as if the fitness landscape did not change at all but was the average of the changing landscapes (Wilke et al. 2006).

**Genetic relatedness between reservoir and naïve hosts.**

A compelling suggestion is that the more closely related the reservoir and the new host are, the greater the chances for a successful spillover (DeFilippis and Villareal 2000). There is a good mechanistic reason to believe that a relationship exists between hosts' phylogenetic distance and the likelihood of viral emergence. If the ability to recognize and infect a host

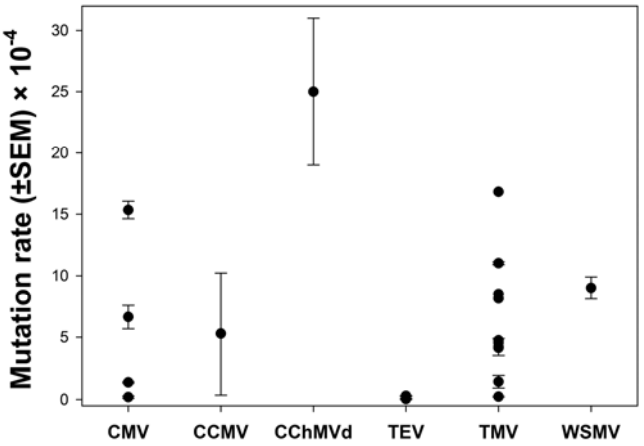


**Fig. 2**. Available mutation rate estimates for different plant RNA viruses and a viroid (per base and per generation). In the case of the five RNA viruses, the mutant spectrum that was generated after inoculation of single plants with an infectious cDNA (i.e, zero initial genetic variability) was characterized and was used to estimate the upper limit of the mutation rate. In the case of the viroid, the actual mutation rate was estimated from the frequency of lethal alleles in the population. The original data can be found in: CMV (Pita et al. 2007; Schneider and Roossinck 2000, 2001), CCMV (Schneider and Roossinck 2000), CChMVd (Gago et al. 2009), TEV (Sanjuán et al. 2009; Tromas and Elena 2010), TMV (Kearney et al. 1999; Malpica et al. 2002; Schneider and Roossinck 2000, 2001), and WSMV (Hall et al. 2001b).

cell is important for cross-species transmission, then related species are more likely to share related vectors, cell receptors, and defense pathways. However, others state that there are no rules to predict the susceptibility of a new host; spillovers have occurred between hosts independently of their relatedness (Holmes and Drummond 2007). Moreover, viral host switches between closely related species (e.g., species within the same genera) may be limited by cross-immunity to related pathogens (Parrish et al. 2008).

In a very recent study, Cronin and associates (2010) evaluated the relative importance that the following four variables had in key epidemiological parameters that determine potential of different species to serve as reservoirs for *Barley yellow dwarf virus* species PAV (BYDV-PAV) and promote spillovers: i) phylogenetic relatedness between host species, ii) differences in physiological phenotype (rapidly growing short-lived leaves and high metabolic rates vs. slow-growing long-lived leaves and low metabolic rates), iii) provenance (exotic vs. naïve), and iv) host lifespan. Host physiological phenotype and not the degree of phylogenetic relatedness was the variable better explaining variation among species in their potential as BYDV-PAV reservoirs. Indeed, differences among host species in the probability of transmission of BYDV-PAV from an infected host to an uninfected feeding vector were only explained by this variable.

**Phase II: the genetics of adaptation to the new host.**

Although we have described the role of mutation and recombination in the context of phase I, we would like to stress here that these factors will also be essential for the success of phase II. Additional beneficial mutations or new genetic combinations would be needed to further ensure adaptation to the new host.

The evolutionary fate of a population in a constant environment depends on the distribution of mutational effects on fitness. This distribution encompasses all possible mutations and can be divided into fractions, beneficial, neutral, deleterious, or lethal. Given the compactness of viral genomes for a well-adapted virus, most mutations are expected to fall into the last two categories. Recently, Carrasco and associates (2007) explored the distribution of single-nucleotide substitution mutational effects for TEV on its natural host, tobacco. Most mutations were strongly deleterious for the virus, with up to 41% being lethal, 36% being significantly deleterious (on average, reducing fitness 41%), 23% having no measurable effect on fitness (i.e., they were neutral), and none being beneficial. However, the distribution of fitness effects on a given genotype is rarely constant across hosts, and the contribution of each category to the overall fitness will vary depending on the overlap between the alternative hosts (Martin and Lenormand 2006). This host-dependence of the distribution of mutational effects may impact the likelihood of adaptation after host switching. For instance, if the host provides new opportunities for the virus, the fraction of beneficial mutations may be increased either by moving the average of the distribution towards more positive values while keeping the shape constant or, alternatively, by increasing the variance without affecting the mean. Characterizing the distribution of mutational effects across a panel of possible alternative hosts varying in genetic relatedness to the natural one is a pending task.

Given the high mutation rate of RNA viruses, mutations may not appear as single events, but genomes may contain multiple hits (Malpica et al. 2002; Tromas and Elena 2010). The way in which mutations interact in determining viral fitness, a concept known as epistasis, conditions whether certain evolutionary pathways are more likely than others (Weinreich et al. 2005). If mutational effects are multiplicative, the shape of the landscape will be smooth, with a single peak emerging from a flat surface. By contrast, the stronger the deviation from multiplicativity, the more fitness peaks of different heights may exist in a landscape. Unfortunately, a direct evaluation of the extent and intensity of epistasis in the genome of plant RNA viruses is not yet available. Only an indirect statistical evaluation exists for TEV, suggesting that, on average, mutations interact in a negative way; that is, the observed effect of two mutations together is lower than expected by multiplying their individual effects (De la Iglesia and Elena 2007). This result is in good agreement with those observed for animal viruses and bacteriophages (Elena et al. 2010). The cause of negative epistasis may be found in the existence of overlapping genes in RNA genomes encoding for multifunctional proteins (Elena et al. 2006).

**A plant is not a test tube: spatial structure and metapopulation dynamics.**

Plant architecture creates a spatially structured environment for plant viruses. This means that the viral population replicating within an infected plant must be considered as a collection of subpopulations, each replicating in different parts, from the arrangement of different tissues within a leaf to individual leaves and, finally, branches. Spatial structure imposes strong conditions on the spread of beneficial mutations that may improve the fitness of an emerging virus on its new host.

Using *Plum pox virus* (PPV) clones labeled with two different colors of fluorescent protein, Dietrich and Maiss (2003) were able to observe that the two populations excluded each other during the colonization of *N. benthamiana* epidermal cells. Only a minority of cells in the contact region between growing foci were doubly infected. Spatial separation reduces the opportunities for competition between genetic variants, thus reducing the efficiency with which natural selection may increase overall population fitness and reducing the opportunity for recombination. Furthermore, this strong spatial structure imposes obstacles to the fixation of beneficial mutations in the whole metapopulation, regardless of the magnitude of their beneficial effect within a particular spatial level. If beneficial mutations appear in cells that are already confined by other infected cells, they will be unable to spread spatially and will, therefore, be invisible to selection. Spatial structure and mutual exclusion also reduces the opportunity for recombination and, thus, of generation of genetic variation.

Jridi and associates (2006) analyzed the population structure of PPV within a single infected *Prunus persica*. They observed that, following the systemic invasion of the host, the virus population differentiated into several subpopulations that were isolated in different branches. These subpopulations subsequently differentiated into other subpopulations, with little to no genetic exchange between distal parts.

One may ask whether this segregation of viral populations into different subpopulations is driven by fitness differences or if the determination of the genotype colonizing distal tissue is a purely stochastic process. In recent years, different groups have evaluated the strength of population bottlenecks during the colonization of distal tissues. Hall and associates (2001a) estimated the effective population size ($N_e$) during systemic colonization of WSMV. In short, they mixed two different strains of WSMV and used the mixture to coinfect wheat seedlings. Then, they determined how many tillers were infected with a single strain versus how many were coinfected. The frequency data were then fitted to a binomial distribution, and it was determined that $N_e$ for systemic colonization was three to five genomes. Sacristán and associates (2003) used a similar coinoculation approach and estimated that, during systemic colonization by TMV, $N_e$ was in the order of units. In a similar experiment that involved 12 genetic markers, Li and Roossinck

(2004) showed that the genetic variance of CMV populations replicating in a single leaf was significantly and reproducibly reduced in systemic leaves, with the number of markers present in the systemic leaves ranging between four and eight, although proper statistical analyses of these data render $N_e$ in the range of 12 to 220. Finally, Monsion and associates (2008) estimated, again using a similar experimental design involving six markers, that $N_e$ for CaMV infecting systemic leaves of *Brassica rapa* was in the range of several hundred genomes. In conclusion, reported $N_e$ estimates differ widely among viruses. Although there will undoubtedly be variation in $N_e$ between viruses, the observed variation may also be attributable, in part, to the experimental setup, in particular, exposure to the virus. In the case of the two large $N_e$ estimates (CMV and CaMV), plants were inoculated with sufficient initial numbers of virions so all plants were infected. The resulting estimates may therefore not be indicative of the minimum $N_e$. Estimating $N_e$ over different doses for a plant virus would be useful (Zwart et al. 2009) and is another pending task.

An additional factor about the spatial spread of genetic variants at high MOI needs to be considered. Complementation between genotypes may slow the rate at which a beneficial mutation spreads in the population (Frank 2001). When many viral genotypes infect the same host cell, the effective ploidy of the genetic system is high, diluting the contribution of each locus to the phenotype and weakening the selective intensity on each locus. Weaker selection allows maintenance of greater genetic diversity in the population, allowing otherwise deleterious alleles to persist for long periods of time.

**Evasion of host defenses.**

Upon entering the new host, a virus needs to cope with plant defenses. Plants have a wide variety of complex responses to viral infection, including innate and acquired nonspecific resistance mechanisms and specific responses (e.g., gene-for-gene, systemic acquired resistance, and RNA silencing). All these forms of resistance have been recently reviewed by others (Loebenstein 2009; Moffett 2009; Ruiz-Ferrer and Voinnet 2009; Truniger and Aranda 2009) and, therefore, we will not discuss the evolutionary solutions that viruses may find to escape from each mechanism. In contrast, we will comment on just one that has represented a revolution in our understanding of the interaction between viral pathogens and plants and that is also an active conserved mechanism in other eukaryotic hosts, virus-induced RNA silencing. Furthermore, we find it particularly interesting because of its properties of memory and sequence-specificity (Ding 2010; Ding and Voinnet 2007; Voinnet 2001). Not surprisingly, soon after the identification of RNA silencing as a defense mechanism, the existence of viral proteins with the capacity of interacting with different components of the silencing pathway, blocking the antiviral response, and enhancing virus accumulation and systemic movement was reported (Anandalakshmi et al. 1998; Brigneti et al. 1998; Kasschau and Carrington 1998; Lucy et al. 2000; Voinnet et al. 1999). The evolutionary implications of these suppressor proteins have not been fully explored yet, but in a recent compensatory-evolution experiment, Torres-Barceló and associates (2010) showed that the TEV suppressor protein HC-Pro may be under strong stabilizing selection, suggesting that it is detrimental for the virus both to reduce and to increase the strength of suppression.

The high mutation rate of plant RNA viruses may itself facilitate evasion of RNA silencing by generating escape mutants at a high frequency. To evaluate the likelihood of generating mutants capable of escaping from the selective pressure imposed by a single siRNA, Lin and associates (2009) inserted a noncoding sequence into the genome of TuMV. This noncoding sequence was targeted by an artificial microRNA expressed by the host plant *N. benthamiana*. As expected, transgenic plants were resistant to TuMV. Each of the 21 nucleotides in the siRNA target sequence was then mutated, and the pathogenicity of each single-nucleotide substitution mutant was evaluated in the transgenic plants. Mutations at six positions in the target sequence rendered viruses with high pathogenicity, mutations at nine other positions only produced a minor increase in pathogenicity. Nonetheless, mutations at any site in the target sequence allowed the mutant virus to replicate to the extent that additional mutations that further increased pathogenicity were generated (Lin et al. 2009). This experiment serves as an example of the ease with which a population of RNA viruses may escape from the surveillance of siRNAs simply by mutation. However, it is worth noting that i) in a more realistic situation multiple siRNAs are produced against the viral genome, and ii) the target sequence encodes a protein, implying that not all changes would be equally tolerated due to functional constraints.

In a recent study, Agudelo-Romero and associates (2008b) found evidence suggesting that adaptation of TEV to *Arabidopsis thaliana* proceeded via downregulation of genes involved in stress responses. These stress genes were all significantly up-regulated in plants infected with the ancestral nonadapted virus (Agudelo-Romero et al. 2008a), suggesting that one outcome of natural selection to optimize viral fitness in a novel host is by making it undetectable to plant defenses.

A possible mechanism by which viruses may directly manipulate the expression of host genes is via sequence homology between virus-derived small RNAs (vsRNA) and host genes. Moissiard and Voinnet (2006) found direct evidence supporting this hypothesis for CaMV. *A. thaliana* and turnip plants infected with CaMV accumulated vsRNA originated from the fold-back structure of the leader sequence located at the 5′ end of the 35S RNA. Using a bioinformatic approach, it was found that the 35S leader sequence has high homology with several *A. thaliana* transcripts, the three most prominent ones being *At1g76950*, *At1g75330*, and *At3g52500* (encoding, respectively, an uncharacterized member of the RCC1 family, an ornithine carbamoyltransferase/transcarbamylase, and an aspartyl protease). It was further experimentally demonstrated that infected plants had reduced expression of these genes. As pointed out by the authors, one question that has not been answered is whether host gene-targeting by CaMV represents a viral strategy to facilitate infection or is collateral damage.

**Conclusions.**

This review ignores one of the main characters in the picture, i.e., the vector. Vectors strongly influence the probability of successful emergence. From an ecological perspective, transmission occurs only between plants on which the vector feeds, thus controlling the intensity of the flow between the reservoir and the new host. From a genetic perspective, vectors impose strong bottlenecks between host-to-host transmissions, during which a large part of the standing variation is lost. Besides, the vector itself can also be involved in tradeoffs in which the adaptation to a new plant host reduces the affinity of the virus for the actual vector and selection may favor the use of new ones.

Most of the material we brought together for this review explores the role of viral evolution in the early stages of emergence. We would like to argue here that the viral genetic variability contained in the reservoir population is the most important genetic determinant of viral emergence. The forest depicted in Figure 1 represents a giant black box because we know little about viruses of wild plant species that probably work as a large reservoir generating spillovers on cultivated plants or between wild species, so there is a whole evolutionary space that

we totally ignore, making it more difficult to predict and prevent emerging plant viral diseases. Natural selection will operate upon this genetic variability to optimize viral fitness during Phase II. After reading the presentation we made above, one may consider that successful emergence, characterized by sustained host-to-host transmission, may be a far more difficult process than expected given the remarkable evolutionary plasticity of RNA viruses. Fitness tradeoffs, pleiotropic fitness effects, strong bottlenecks at different levels, an excess of deleterious mutations, spatial constraints, and fragmented host populations will limit the chances for new viruses to emerge. Therefore, the emergent viruses that we are witnessing at this time may represent just the few lucky cases that have been able to surmount all these limitations.

## ACKNOWLEDGMENTS

This work has been supported by grants BFU2009-06993 (Ministerio de Ciencia e Innovación, Madrid, Spain), RGP2008/12 (Human Frontiers Science Organization), ACOMP2010-089 and PROMETEO2010-019 (Generalitat Valenciana).

## LITERATURE CITED

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

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

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

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., and Vance, V. B. 1998. A viral suppressor of gene silencing in plants. Proc. Natl. Acad. Sci. U.S.A. 95:13079-13084.

Awadalla, P. 2003. The evolutionary genomics of pathogen recombination. Nat. Rev. Genet. 4:50-60.

Brigneti, G., Voinnet, O., Li, W. X., Ji, L. H., Ding, S. W., and Baulcombe, D. C. 1998. Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. EMBO (Eur. Mol. Biol. Organ.) J. 17:6739-6746.

Bruyere, A., Wantroba, M., Flasinski, S., Dzianott, A., and Buyarski, J. J. 2000. Frequent homologous recombination events between molecules of one RNA component in a multipartite RNA virus. J. Virol. 74:4214-4219.

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

Chare, E. R., and Holmes, E. C. 2006. A phylogenetic survey of recombination frequency in plant RNA viruses. Arch. Virol. 151:933-946.

Chare, E. R., Gould, E. A., and Holmes, E. C. 2003. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. J. Gen. Virol. 85:3149-3157.

Cleaveland, S., Haydon, D. T., and Taylor, L. 2007. Overviews of pathogen emergence: Which pathogens emerge, when and why? Curr. Top. Microbiol. Immunol. 315:85-111.

Codoñer, F. M., and Elena, S. F. 2008. The promiscuous evolutionary history of the family *Bromoviridae*. J. Gen. Virol. 89:1739-1747.

Cronin, J. P., Welsh, M. E., Dekkers, M. G., Abercrombie, S. T., and Mitchell, C. E. 2010. Host physiological phenotype explains pathogen reservoir potential. Ecol. Lett. 13:1221-1232.

DeFilippis, V. R., and Villareal, L. P. 2000. An introduction to the evolutionary ecology of viruses. Pages 126-208 in: Viral Ecology. C. J. Hurst, ed. Academic Press, New York.

De la Iglesia, F., and Elena, S. F. 2007. Fitness declines in *Tobacco etch virus* upon serial bottleneck transfers. J. Virol. 10:4941-4947.

Dietrich, C., and Maiss, E. 2003. Fluorescent labeling reveals spatial separation of potyvirus populations in mixed infected *Nicotiana bethamiana* plants. J. Gen. Virol. 84:2871-2876.

Ding, S. W. 2010. RNA-based antiviral immunity. Nat. Rev. Immunol. 10:632-644.

Ding, S. W., and Voinnet, O. 2007. Antiviral immunity directed by small RNAs. Cell 130:413-426.

Elena, S. F., and Sanjuán, R. 2008. Virus evolution: Insights from an experimental approach. Annu. Rev. Ecol. Evol. Syst. 38:27-52.

Elena, S. F., Carrasco, P., Daròs, J. A., and Sanjuán, R. 2006. Mechanisms of genetic robustness in RNA viruses. EMBO (Eur. Mol. Biol. Organ.) Rep. 7:168-173.

Elena, S. F., Agudelo-Romero, P., and Lalic, J. 2009. The evolution of viruses in multi-host fitness landscapes. Open Virol. J. 3:1-6.

Elena, S. F., Solé, R. V., and Sardanyés, J. 2010. Simple genomes, complex interactions: Epistasis in RNA virus. Chaos 20:026106

Escriu, F., Fraile, A., and García-Arenal, F. 2000. Evolution of virulence in natural populations of the satellite RNA of *Cucumber mosaic virus*. Phytophatology 90:480-495.

Frank, S. A. 2001. Multiplicity of infection and the evolution of hybrid incompatibilities in segmented viruses. Heredity 87:522-529.

Froissart, R., Roze, D., Uzest, M., Galibiert, L., Blanc, S., and Michalakis, Y. 2005. Recombination every day: Abundant recombination in a virus during a single multi-cellular host infection. PLoS Biol. 3:e89. Published online.

Gago, S., Elena, S. F., Flores, R., and Sanjuán, R. 2009. Extremely high mutation rate of a hammerhead viroid. Science 323:1308.

García-Arenal, F., Fraile, A., and Malpica, J. M. 2001. Variability and genetic structure of plant virus populations. Annu. Rev. Phytopathol. 39:157-186.

González-Jara, P., Fraile, A., Cantó, T., and García-Arenal, F. 2009. The multiplicity of infection of a plant virus varies during colonization of its eukaryotic host. J. Virol. 83:7487-7494.

Gutiérrez, S., Yvon, M., Thébaud, G., Monsion, B., Michalakis, Y., and Blanc, S. 2010. Dynamics of the multiplicity of cellular infection in a plant virus. PLoS Pathog. 6:e1001113. Published online.

Hall, J. S., French, R., Hein, G. L., Morris, T. J., and Stenger, D. C. 2001a. Three distinct mechanisms facilitate genetic isolation of sympatric *Wheat streak mosaic virus* lineages. Virology 282:230-236.

Hall, J. S., French, R., Morris, T. J., and Stenger, D. C. 2001b. Structure and temporal dynamics of populations within *Wheat streak mosaic virus* isolates. J. Virol. 75:10231-10243.

Holmes, E. C. 2008. The evolutionary history and phylogeography of human viruses. Annu. Rev. Microbiol. 62:307-328.

Holmes, E. C., and Drummond, A. J. 2007. The evolutionary genetics of viral emergence. Curr. Top. Microbiol. Immunol. 315:51-66.

Jenner, C. E., Wang, X., Ponz, F., and Walsh, J. A. 2002. A fitness cost for *Turnip mosaic virus* to overcome host resistance. Virus Res. 86:1-6

Jridi, C., Martin, J. F., Mareie-Jeanne, V., Labonne, G., and Blanc, S. 2006. Distinct viral populations differentiate and evolve independently in a single perennial host plant. J. Virol. 80:2349-2357.

Kasschau, K. D., and Carrington, J. C. 1998. A counterdefensive strategy of plant viruses: Suppression of posttranscriptional gene silencing. Cell 95:461-470.

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

Kearney, C. M., Thomson, M. J., and 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.

Li, H., and Roossinck, M. J. 2004. Genetic bottlenecks reduce population variation in an experimental RNA virus population. J. Virol. 78:10582-10587.

Lin, S. S., Wu, H. W., Elena, S. F., Chen, K. C., Niu, Q. W., Yeh, S. D., Chen, C. C., and Chua, N. H. 2009. Molecular evolution of a viral noncoding sequence under the selective pressure of amiRNA-mediated silencing. PLoS Pathog. 5:e1000312. Published online.

Loebenstein, G. 2009. Local lesions and induced resistance. Adv. Virus Res. 75:73-117.

Lucy, A. P., Guo, H. S., Li, W. X., and Ding, S. W. 2000. Suppression of post-transcriptional gene silencing by a plant viral protein localized in the nucleus. EMBO (Eur. Mol. Biol. Organ.) J. 19:1672-1680.

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

Martin, G., and Lenormand, T. 2006. The fitness effect of mutations across environments: A survey in light of fitness landscapes models. Evolution 60:2413-2427.

Miyashita, S., and Kishino, H. 2010. Estimation of the size of genetic bottlenecks in cell-to-cell movement of *Soil-borne wheat mosaic virus* and the possible role of the bottlenecks in speeding up selection of variations in *trans*-acting genes or elements. J. Virol. 84:1828-1837.

Moffett, P. 2009. Mechanisms of recognition in dominant R gene mediated resistance. Adv. Virus Res. 75:1-33.

Monsion, B., Froissart, R., Michalakis, Y., and Blanc, S. 2008. Large bottleneck size in *Cauliflower mosaic virus* populations during host plant colonization. PLoS Pathog. 4:e1000174. Published online.

Moissiard, G., and Voinnet, O. 2006. RNA silencing of host transcripts by *Cauliflower mosaic virus* requires coordinated action of the four *Arabi-*

*dopsis* Dicer-like proteins. Proc. Natl. Acad. Sci. U.S.A. 103:19593-19598.
Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E. C., Burke, D. S., Calisher, C. H., Laughlin, C. A., Saif, L. J., and Daszak, P. 2008. Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev. 72:457-470.
Pita, J. S., De Miranda, J. R., Schneider, W. L., and Roossinck, M. J. 2007. Environment determines fidelity for an RNA virus replicase. J. Virol. 81:9072-9077.
Ruiz-Ferrer, V., and Voinnet, O. 2009. Roles of plant small RNAs in biotic stress responses. Annu. Rev. Plant Biol. 60:485-510.
Sacristán, S., Malpica, J. M., Fraile, A., and García-Arenal, F. 2003. Estimation of population bottlenecks during systemic movement of *Tobacco mosaic virus* in tobacco plants. J. Virol. 77:9906-9911.
Sanjuán, R., Agudelo-Romero, P., and Elena, S. F. 2009. Upper-limit mutation rate estimation for a plant RNA virus. Biol. Lett. 5:394-396.
Schneider, W. L., and 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.
Schneider, W. L., and Roossinck, M. J. 2001. Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. J. Virol. 75:6566-6571.
Thresh, J. M. 2006. Plant virus epidemiology: The concept of host genetic vulnerability. Adv. Virus Res. 67:89-125.
Torres-Barceló, C., Daròs, J. A., and Elena, S. F. 2010. Compensatory molecular evolution of HC-Pro, an RNA-silencing suppressor from a plant RNA virus. Mol. Biol. Evol. 27:543-551.
Tromas, N., and Elena, S. F. 2010. The rate and spectrum of spontaneous mutations in a plant RNA virus. Genetics 185:983-989.
Truniger, V., and Aranda, M. A. 2009. Recessive resistance to plant viruses. Adv. Virus Res. 75:119-159.
Urbanowicz, A., Alejska, M., Formanowicz, P., Blazewicz, J., Figlerowicz, M., and Bujarski, J. J. 2005. Homologous crossover among molecules of *Brome mosaic bromovirus* RNA1 or RNA2 segments *in vivo*. J. Virol. 79:5732-5742.
Voinnet, O. 2001. RNA silencing as a plant immune system against viruses. Trends Genet. 17:449-459.
Voinnet, O., Pinto, Y. M., and Baulcombe, D. C. 1999. Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses of plants. Proc. Natl. Acad. Sci. U.S.A. 24:14147-14152.
Wallis, C. M., Stone, A. L., Sherman, D. J., Damsteegt, V. D., Gildow, F. E., and Schneider, W. L. 2007. Adaptation of *Plum pox virus* to a herbaceous host (*Pisum sativum*) following serial passages. J. Gen. Virol. 88:2839-2845.
Weinreich, D. M., Watson, R. A., and Chao, L. 2005. Sign epistasis and genetic constraints on evolutionary trajectories. Evolution 59:1165-1174.
Wilke, C. O., Forster, R., and Novella, I. S. 2006. Quasispecies in time-dependent environments. Curr. Top. Microbiol. Immunol. 299:33-50.
Woolhouse, M. E. J., and Dye, C. 2001. Population biology of emerging and reemerging pathogens—Preface. Phil. Trans. R. Soc. Lond. B 356:981-982.
Zwart M. P., Hemerik L, Cory, J. S., De Visser, J. A. G. M., Bianchi, F., Van Oers, M. M., Vlak, J. M., Hoekstra, R. F., and Van der Werf, W. 2009. An experimental test of the independent action hypothesis in virus-insect pathosystems. Proc. R. Soc. B 276:2233-2242.