Mechanisms of genetic robustness in RNA viruses

Santiago F. Elena*, Purificación Carrasco, José-Antonio Daròs & Rafael Sanjuán
Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV), Valencia, Spain

Two key features of RNA viruses are their compacted genomes and their high mutation rate. Accordingly, deleterious mutations are common and have an enormous impact on viral fitness. In their multicellular hosts, robustness can be achieved by genomic redundancy, including gene duplication, diploidy, alternative metabolic pathways and biochemical buffering mechanisms. However, here we review evidence suggesting that during RNA virus evolution, alternative robustness mechanisms may have been selected. After briefly describing how genetic robustness can be quantified, we discuss mechanisms of intrinsic robustness arising as consequences of RNA-genome architecture, replication peculiarities and quasi-species population dynamics. These intrinsic robustness mechanisms operate efficiently at the population level, despite the mutational sensitivity shown by individual genomes. Finally, we discuss the possibility that viruses might exploit cellular buffering mechanisms for their own benefit, producing a sort of extrinsic robustness.

Keywords: fitness; deleterious mutations; quasi-species; genetic robustness; virus evolution


Introduction

RNA viruses have the highest mutation rate among living species (that is, between 10⁻³ and 10⁻² errors per nucleotide and replication cycle), very small and compacted genomes, short generation times and extremely large populations (Domingo & Holland, 1997). This might be beneficial in the long-term, as it allows viral populations to quickly explore genotypic space and find beneficial mutations. However, it is clearly detrimental in the short-term as most mutations have deleterious fitness effects. The balance between the continuous generation of mutants and the action of selection leads to a dynamic population structure, known as ‘quasi-species’ (Domingo & Holland, 1997).

In recent years, the interest of evolutionary biologists in the mechanisms, consequences and evolution of genetic robustness has been revitalized by new and powerful techniques that allow the tracking and manipulation of genotypes (de Visser et al, 2003). Robustness is defined as a reduced sensitivity to perturbations affecting phenotypic expression. If perturbations are inheritable, then we talk about genetic robustness; if they are not (for example, changes in physical and chemical parameters, or developmental noise), then we talk about environmental robustness. Robustness should occur when there are several copies of a single gene, when several genes contribute to the same function or through biochemical buffering mechanisms. This includes gene duplication, polyploidy, alternative metabolic pathways or chaperone proteins. As illustrated in Fig 1A, a lack of robustness is expected in haploid genomes that have no duplications, overlapping gene functions, repair systems and arepleiotropic. A small number of mutations can produce a strong effect, but as mutations accumulate, they affect the same function with increasing probability and, thus, their marginal contribution to fitness diminishes. Hence, the observed fitness is above the expected multiplicative value or, in other words, epistasis is antagonistic (Wolf et al, 2000). By contrast, in the presence of redundancy and buffering mechanisms, the fitness of genomes is only mildly affected; however, as the mutation load increases, these mechanisms ultimately collapse. Fitness will therefore be lower than the expected multiplicative value, which means that there will be synergistic epistasis (Fig 1B).

In principle, genetic robustness might evolve for one of the following reasons. First, as long as robustness has a heritable basis, shows variability among individuals and affects the probability of survival, it can be a target for selection and evolutionary optimization (Wilke & Adami, 2003). The selection pressure for increasing robustness depends on the occurrence of mutations. The more frequent mutations are, the more efficient selection will be at promoting the evolution of robustness. Second, it might evolve because buffering is a necessary consequence of character adaptation; that is, robustness is a side-effect of stabilizing selection acting on different traits (Meiklejohn & Hartl, 2002). Third, given that environmental fluctuations often have a strong impact on fitness, selection would efficiently favour mechanisms of environmental robustness. On the basis of theoretical arguments and RNA folding simulations, some authors have predicted that genetic robustness should be intrinsically correlated to environmental robustness and, thus, that the former could evolve as a correlated response to selection favouring the latter (Ancel & Fontana, 2000; Wagner et al, 1997). This is an appealing hypothesis because, during their life cycle, RNA viruses must cope not only with the deleterious effect of mutations but also with dramatic and fast fluctuations in their environments such as alternating among host species, tissue- and organ-specific microenvironments or the presence of antiviral agents.
Robustness in RNA viruses
S.F. Elena et al

Measuring genetic robustness

Although it might be interesting to measure mutational robustness in specific biochemical pathways, gene expression patterns or developmental plans, it is more informative, from an evolutionary standpoint, to determine the effects of mutations on fitness. To quantify the fitness effects of a given mutation, i.e., population geneticists make use of the selection coefficient $s = W_i - 1$, in which $W_i$ is the relative fitness of the mutant; that is, the number of descendants per wild-type descendant. Most mutations are deleterious and thus the average selection coefficient, $s$, is negative. Whenever information about the average fitness effects of mutations is available, $s$ can be seen as a measure of mutational sensitivity and, henceforth, as an inverse approximation to robustness.

A simple experimental approach to estimate $s$ is the Bateman–Mukai mutation-accumulation (MA) method (Mukai, 1964). Such experiments consist of the parallel accumulation of spontaneous mutations in which the effect of natural selection is minimized—for example, by population bottlenecks or by endogamic crosses. The main limitation of this kind of experiment is that highly unviable or lethal mutations cannot accumulate, and therefore the average effects of mutations tend to be underestimated. This drawback can be minimized by a mutagenesis approach, using radiation, chemical treatments, transposons, proofreading-defective polymerases or site-directed mutagenesis.

RNA virus hypersensitivity to deleterious mutations

During the past decade, virologists have established that RNA viruses are very sensitive to the effect of deleterious mutations. In a pioneering study, Chao (1990) performed an MA experiment by imposing dramatic population bottlenecks on the phage 66. After the bottlenecks took effect, the average fitness of 66 was reduced by 22%. This experiment was followed by others under similar demographic circumstances: vesicular stomatitis (VSV), foot and mouth disease (FMDV) and human immunodeficiency type 1 (HIV-1) viruses suffered fitness losses of 38%, 35% and 82%, respectively (Duarte et al., 1992; Escarmís et al., 1996; Yuste et al., 1999), and bacteriophage MS2 lost as much as 16% fitness in a single bottleneck event (de la Peña et al., 2000). The rate of log-fitness decay equals $U \times s$ (where $U$ is the genomic mutation rate) and, therefore, high fitness losses can be explained by a high $U$-value. However, when Escarmís et al. (1996) obtained consensus sequences for the bottlenecked FMDV populations, only a few mutations were fixed in each lineage, suggesting that each one imposed a strong fitness burden. A similar conclusion was obtained for HIV-1 (Yuste et al., 2000). Indeed, there are several strategies for disentangling the contributions of $U$ and $s$ to the rate of fitness decline, such as computing the among-lines variance in MA experiments, sequencing the MA lines, or undertaking a mutagenesis experiment. Using site-directed mutagenesis, Sanjuán et al. (2004a) showed that up to 40% of random single-nucleotide substitutions in VSV were lethal and that, for non-lethal mutations, fitness reductions were as high as 46% (supplementary Table S1 online).

After an extensive review of the literature on mutational effects (supplementary Table S1 online), the median selection coefficient against single mutations across different RNA viruses has been calculated to $~10.8%$. This high sensitivity to mutation contrasts sharply with the estimates obtained for DNA organisms, in which many mutations are believed to be neutral (Conant & Wagner, 2004; Davies et al., 1999; Steinmetz et al., 2002) and a review of available data on mutational fitness effects in these organisms yields a median deleterious mutational effect across species of $~1.7%$ per generation (supplementary Table S2 online). Therefore, the data support the idea that RNA viruses are genetically unprotected against mutation, possibly due to the compactness of their genomes, with frequently overlapping functional regions and few alternative pathways.

Although $s$ provides important information about genetic robustness, it might not reflect the complete picture because when several mutations are present simultaneously in the same genotype, they might interact to determine fitness. As we hypothesized above, synergistic epistasis would be diagnostic for robustness mechanisms, whereas antagonistic epistasis would be characteristic of hypersensitive genomes. Although reports of synergistic epistasis in eukaryotes exist (de Visser & Lenski, 1998; Elena & Lenski, 1997), among RNA viruses, no evidence for epistasis was obtained from MA experiments with phage MS2 (de la Peña et al., 2000), FMDV (Elena, 1999) and poliovirus (Burch & Chao, 2004), although in these studies the number and type of the mutations accumulated was unknown and the statistical power was probably too low. However, more recent and exhaustive MA and site-directed mutagenesis experiments carried out with 66 (Burch & Chao, 2004), VSV (Sanjuán et al., 2004b) and HIV-1 (Bonhoeffer et al., 2004) have revealed a tendency towards antagonistic epistasis.

For loci to be epistatic, they must be involved in the expression of the same trait. Whether this produces antagonistic or synergistic epistasis depends on the way these loci participate in the expression of the character. If the contribution of several loci is necessary, then epistasis will be antagonistic, as the effect of a mutation at one locus is the same as mutations at several loci; this is the case when the loci involved belong to the same gene. Antagonistic epistasis should
therefore be characteristic of small and compact genomes. Finally, in RNA genomes, there is a strong selection pressure for maintaining secondary structures, and compensatory mutations can occur relatively often (Wilke et al., 2003). Both features might explain the tendency of RNA viruses to display antagonistic epistasis.

Population mechanisms of intrinsic viral robustness

RNA viruses show hypersensitivity to mutational effects, but the question remains: how do they face deleterious mutations? Although it might sound paradoxical, individual mutational hypersensitivity produces robustness at the population level. The efficiency of natural selection to purge deleterious alleles from a population, or more generally, the relative contribution of selection and genetic drift to evolutionary change, depends on the product $N \times s$ (where $N$ is the population size). The efficiency of natural selection increases at high $N$ provided that ecological resources are limited. In this scenario, a form of soft selection operates through the differential survival and associated competition among individuals. This is likely to be the case in RNA viruses. Usually, enormous values of $N$ are reached even in a single infected host and many more viral particles are produced than available cells. The contribution of large $N$ acts in positive synergy with the above-mentioned large $s$, making selection remarkably efficient. Indeed, Krakauer & Plotkin (2002) showed that lineages existing as very large $N$-values might become more robust at the population level, whereas individual sensitivity to mutation is increased. This anti-redundant strategy relies on the efficient elimination of mutated genotypes and the preservation of the non-mutated genotype.

Population genetics theory predicts that the optimal evolutionary mutation rates are largely determined by $s$ (Johnson & Barton, 2002; Orr, 2000) and thus, the mutational hypersensitivity shown by RNA viruses might allow them to tolerate particularly high mutation rates. This interpretation provides an evolutionary basis for the characteristically high mutation rates observed in RNA viruses but apparently fails to account for the concomitantly high variability of these systems. However, in good agreement with the anti-redundancy strategy, increasing evidence suggests that much of this variation is rapidly purged from populations and that only a few beneficial changes are responsible for adaptation (Novella et al., 1999; Rodriguez et al., 2000).

Opposite to the anti-redundancy strategy is the idea of ‘survival of the flattest’, originally postulated by Schuster & Swetina (1988). Standard population genetics theory has shown that, under some simplifying assumptions, the average fitness of a population equals $\exp(-U)$ (Kimura & Maruyama, 1966). However, when neutral and back mutations are considered, the population average equilibrium fitness also depends on the geometry of the fitness landscape, which, at a first instance, can be described by $s$ (Wilke & Adami, 2003). This being the case, another selective pressure comes into play at high values of $U$ that pushes populations towards regions of the landscape where the density of neutral mutants is higher (Wilke, 2001). As a consequence, the whole population evolves increased robustness against mutations. This phenomenon of selection for robustness at high values of $U$ can be understood as a pressure for populations to occupy either highly connected areas on a neutral network or broad rather than narrow fitness peaks; populations occupying high but narrow peaks are easily mutated to have low fitness, whereas populations living in low but broad peaks are less susceptible to mutations (Fig 2). Despite the fact that the effect of survival of the flattest has been observed in silico (Wilke et al., 2001), and has been supported by some in vitro observations (Burch & Chao, 2000), definitive experimental proof for its relevance is still missing.

A common misconception is that most RNA viruses are haploid. However, it is more convenient to consider viruses as $n$-ploid systems, $n$ being variable during the infectious cycle. At initial stages, the multiplicity of infection (infectious particles per available cell) is low and, consequently, viruses are effectively haploid. However, as the host is invaded, high multiplicities of infection ensure frequent co-infection events and increasing ploidy. An immediate consequence of polyploidy is genetic complementation. If all co-existing genomes in an infected cell contribute to the same protein pool, and provided that viral proteins diffuse freely in the cytoplasm, mutant genomes can use proteins in which they are deficient as long as they are produced by a co-existing genome. In agreement with this idea, the maintenance of lethal mutations by complementation has been a common observation in animal, plant and bacterial viruses (Cicin-Sain et al., 2005; Moreno et al., 1997; Roux et al., 1991). For example, recent experiments have shown that whenever the probability of co-infection is maximized, low-fitness mutant VSV genotypes can co-exist with the wild-type virus in a frequency-dependent manner (Wilke et al., 2004).

Another population robustness mechanism that might be important for RNA viruses is sex, as it results in not only recombination between homologous molecules, but also the segregation of segments in a multipartite genome. Sex recreates mutation-free genotypes and helps to keep the average population fitness high.
(Otto & Lenormand, 2002). Both forms of sex are common among RNA viruses (Lai, 1992). The fitness advantages associated with recombination and segregation are amply illustrated. Fernández-Cuartero et al (1994) found that a recombinant RNA 3 created from cucumber mosaic and tomato aspermy viruses outcompeted both parental species. Chao et al (1997) found that populations of the tripartite φ6, which were previously subjected to MA (Chao, 1990), recovered fitness through compensatory mutations. However, hybrids produced by crossing the populations recovered fitness to a greater extent, illustrating the positive effect of segregation.

**Mechanisms of extrinsic robustness: cellular chaperones**

Heat-shock proteins (HSPs) are a set of highly conserved proteins, the expression of which is induced by a variety of cellular stresses; however, most of them also have essential functions under normal physiological conditions. Chaperones are a specific subset of the HSPs (Hartl et al., 1994). They belong to structurally unrelated protein families that share the ability to recognize and bind to aberrant protein conformations, such as solvent-exposed hydrophobic domains. Under normal physiological conditions, chaperones function in different ways: by preventing nascent polypeptide chains from forming aberrant associations before the remainder of the protein has emerged from the ribosome; by assisting the entrance of newly synthesized polypeptides into organelles; or by masking the exposed surfaces of oligomeric proteins before they assemble with their partners. Under stressful conditions, chaperones prevent misfolding and aggregation, or can even restore the proper conformation of proteins and actively disaggregate damaged ones. Mutations may lead to the destabilization of proteins and increase their tendency to misfold and aggregate. Chaperones can keep mutationally altered proteins functional and thereby buffer detrimental mutations (Fares et al., 2002).

It is well known that viral infections induce the cellular stress response. Indeed, increases in Hsp70 chaperone levels after viral infection have been widely observed (Mayer, 2005). Furthermore, during viral infection of animals, host chaperones that bind to viral particles can act as flags on the cellular surface, triggering the immune response (Oglesbee et al., 2002). However, is it possible that viruses also use chaperones for their own benefit to buffer mutational effects? It has been shown that most viruses need cellular chaperones during their life cycle both to solve their own protein-folding problems as well as to interfere with cellular processes such as signal transduction (Mayer, 2005). Table 1 provides examples of the interaction between cellular chaperones and viral factors at different stages of the infectious cycle. The fact that mutant misfolded viral proteins elicit the expression of chaperones has been illustrated by Jockusch et al (2001). They showed that the presence of tobacco mosaic virus mutant misfolded coat proteins (CPs) triggers the overexpression of Hsp70, with the intensity of the response proportional to the number of mutant CPs. By contrast, the presence of wild-type CPs does not elicit such a response. Moreover, perhaps the most convincing evidence of the importance of chaperones during the virus life cycle is the existence of virus-encoded HSP-like proteins (Table 1, beet yellow closterovirus Hsp70h). Taken together, these observations indicate that viruses might benefit from cellular chaperones as an extrinsic robustness mechanism.

### Table 1 | Phases of the viral infectious cycle in which chaperones have a role

<table>
<thead>
<tr>
<th>Phase of the infectious cycle</th>
<th>Examples</th>
<th>HSP involved</th>
<th>Viral factor</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell entry</td>
<td>Rotavirus</td>
<td>Hsc70</td>
<td>VP5 coat protein</td>
<td>Guerrero et al (2002)</td>
</tr>
<tr>
<td></td>
<td>Human T lymphotropic type 1</td>
<td>Hsc70</td>
<td>gp46</td>
<td>Fang et al (1999)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hsc70 + Bag3</td>
<td>Penton (fibre protein)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hsp70, Hdj1, Hdj2</td>
<td>DNA helicase E1</td>
<td>Liu et al (1998)</td>
</tr>
<tr>
<td>Gene expression</td>
<td>Flock house</td>
<td>Hsp60</td>
<td>RNA-replication complex assembly</td>
<td>Kampmueller &amp; Miller (2005)</td>
</tr>
<tr>
<td></td>
<td>Tobacco mosaic</td>
<td>Hsp101</td>
<td>S′ leader W sequence</td>
<td>Gallic (2002)</td>
</tr>
<tr>
<td></td>
<td>Polyomavirus</td>
<td>Hsp70</td>
<td>VP1, VP2, VP3</td>
<td>Cripe et al (1995)</td>
</tr>
</tbody>
</table>

Hsc, heat-shock cognate; Hsp, heat-shock protein; VP, viral protein.

**Concluding remarks**

Far from being passive victims of their error-prone replication, RNA viruses cope with the deleterious effects associated with an excess of mutations. In sharp contrast to their cellular hosts, these buffering mechanisms are generally not encoded in their genomes, but are rather a consequence of their population dynamics and parasitic lifestyle. Unravelling the mechanisms of robustness in RNA viruses is a promising research avenue for understanding their survival and evolution.

**Supplementary information** is available at EMBO reports online (http://www.emboreports.org).

**ACKNOWLEDGEMENTS**

S.F.E. is supported by grants from the Spanish Ministerio de Educación y Ciencia (MEC-FEDER), the Generalitat Valenciana, and the EMBO Young Investigator Program. J.-A.D. is supported by grants from MEC-FEDER and the Generalitat Valenciana. R.S. is contracted under the Consejo Superior de Investigaciones Científicas I3P postdoctoral programme.
REFERENCES
Chao L (1990) Fitness of RNA virus decreased by Muller’s ratchet. Nature 348: 454–455
Moreno IM, Malpica JM, Rodríguez-Cerezo E, García-Arenal F (1997) A mutation in tomato aspermy cucumovirus that abolishes cell-to-cell movement is maintained to high levels in the viral population by complementation. J Virol 71: 9517–9527

Robustness in RNA viruses
S.F. Elena et al
Robustness in RNA viruses
S.F. Elena et al


S.F. Elena (who is a member of the EMBO Young Investigator Programme), P. Carrasco, J.-A. Darós & R. Sanjuán