Cell PRESS

Pacing a small cage: mutation and RNA viruses

Robert Belshaw[1], Andy Gardner[2,3], Andrew Rambaut[2] and Oliver G. Pybus[1]

[1] Department of Zoology, University of Oxford, Oxford OX1 3PS, UK
[2] Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK
[3] St John's College, University of Oxford, Oxford OX1 3JP, UK

RNA viruses have an extremely high mutation rate, and we argue that the most plausible explanation for this is a trade-off with replication speed. We suggest that research into further increasing this mutation rate artificially as an antiviral treatment requires a theoretical reevaluation, especially relating to the so-called error threshold. The main evolutionary consequence of a high mutation rate appears to have been to restrict RNA viruses to a small genome; they thus rapidly exploit a limited array of possibilities. Investigating this constraint to their evolution, and how it is occasionally overcome, promises to be fruitful. We explain the many terms used in investigating RNA viral evolution and highlight the specific experimental and comparative work that needs to be done.

The mutation rate of RNA viruses

The single most important feature of RNA viruses is their high mutation rate. Estimates of this rate fall between 0.4 and 1.1 nucleotide errors per genome per round of replication (excluding some outlying retroviruses, which we discuss later) [1]. These mutation rates affect every aspect of virus biology and are at least a hundred-fold higher than those estimated for DNA viruses and other DNA microbes [1]. The difference in rates appears to result from the lack of proofreading by the RNA-dependent polymerases of RNA viruses [2]: the DNA-dependent replicative polymerases of many other organisms have similar misincorporation rates, on the order of 10^{-4} to 10^{-5} per base per round of replication, but the error rate is then reduced to 10^{-5} to 10^{-7} by subsequent proofreading [3].

At its simplest, an RNA virus is a single RNA molecule within a protein shell that enters a host cell and is translated, with the resulting proteins initiating viral replication and leading to the production of many more viral particles by the host cell. At present, there are complete genome sequences for ~500 species of such single-stranded (ss) positive-sense viruses (http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/viruses.html), which infect animals, plants and bacteria (and include poliovirus, foot-and-mouth disease virus [FMDV] and the 'common cold' rhinoviruses). The other RNA virus groups, each with ~100 completely sequenced genomes, are the ss negative-sense viruses (in which the genomic RNA is copied to form mRNA immediately following entry into the cell, and which include influenza, measles, mumps and rabies viruses); the double-stranded viruses (which include the important diarrhoea-causing rotaviruses and the veterinary pest bluetongue); and the retrotranscribing viruses, which convert RNA to DNA as part of their replication cycle (including hepatitis B virus and the retrovirus HIV).

RNA viruses feature prominently in the list of the most serious infectious diseases. The second and sixth biggest killers worldwide are RNA viruses (HIV and measles, respectively), and several RNA viruses contribute significantly to the first and third biggest killers: lower respiratory infections and diarrhoea, respectively [4]. By contrast, although a similar number of DNA viral species are known (whose replication involves DNA to DNA copying), none of these appear in the top 30 of this list.

Here we ask why RNA viruses have such high mutation rates and what are the consequences of these, both for the virus population and their human hosts. We suggest that the most likely cause, and one that is amenable to further empirical study, is a trade-off between replication fidelity and replication speed. An important consequence, which is now exploited therapeutically, is that RNA viruses might be particularly susceptible to further elevation of their mutation rate; however, we argue that explanations for this involving a so-called error threshold might be misleading. The high mutation rate of RNA viruses affects their importance as human pathogens: it can facilitate rapid escape from adaptive immune responses and from drug treatment.

Glossary

Error catastrophe: in the quasispecies model, the loss of the fittest sequence owing to competition with mutated sequences. The term is also used to describe an accompanying loss of the consensus sequence for the population, which then drifts randomly through sequence space.
Error threshold: in the quasispecies model, the point at which fidelity of sequence replication is too low to prevent the error catastrophe (Box 1).
Lethal mutagenesis: the reduction in growth rate of a virus achieved by artificially increasing its mutation rate.
Mutational robustness: the ability of a genotype to sustain mutations without affecting its phenotype.
Quasispecies: a mathematical representation of population growth where the population is divided into categories that are defined by, and linked via, their number of (deleterious) mutations. There is no necessary conflict between these models and those of traditional population genetics: the quasispecies model can be interpreted in terms of mutation–selection balance [46]. Also, the term refers to a chemical rather than a biological definition of a species (i.e. 'almost' a single type or 'species' of molecule). Importantly, and perhaps controversially, mutational linkage leads fitness to being seen as an attribute of the population rather than of an individual virus.
Survival of the flattest: under a high mutation rate, the competitive advantage of a genotype with lower replication rate but higher mutational robustness.

Corresponding author: Belshaw, R. (robert.belshaw@zoo.ox.ac.uk).

0169-5347/$ – see front matter doi:10.1016/j.tree.2007.11.010 Available online 4 March 2008

However, a less-studied consequence is that high mutation rates are likely to constrain viral evolutionary change. Understanding the nature of this constraint, and perhaps more importantly what enables RNA viruses to occasionally escape it, will lead to valuable insights.

Causes of a high mutation rate

Hypotheses to explain why RNA viruses have such a high mutation rate fall into three main categories.

Life history

Many RNA viruses infect hosts that have adaptive immune systems (defences which learn to recognize and destroy invading pathogens) [5,6]. A high mutation rate might be an adaptation to this mode of life because viruses would be more likely to generate mutations enabling them to remain undetected by the immune system of the host for longer. Such mutations provide a striking exception to the general argument that, because most mutations are harmful, natural selection will cause the mutation rate to decrease to the point where it is balanced by the prohibitive metabolic cost of perfect replication fidelity [7].

Although models for RNA viruses have been proposed that balance beneficial mutations against those with harmful effects [8], there is no good correlation between mutation rate and life history [9]. The mutation rates of RNA viruses that attack bacteria, and hence do not face an adaptive immune response, are also high [1,10], and many RNA viruses do not use mutation as a means of evading the adaptive immune response; for example, many use a 'hit-and-run' strategy of being transmitted from the host before the adaptive immune system can respond [5]. Thus, the high mutation rate of RNA viruses cannot readily be ascribed to their life history.

Evolutionary constraint

The high mutation rate could be an evolutionary constraint for RNA viruses. In other words, the high error rate of RNA-dependent polymerases might be something that RNA viruses have simply been unable to improve upon [1]. Also, unlike DNA viruses, RNA viruses do not have the option of using host polymerases for their replication (there are no RNA-dependent polymerases in the host). This is another tempting explanation but, for the reasons described below, it is most likely incorrect as well.

The natural variation in mutation rates among RNA viruses is incompatible with such a constraint. A low rate of mutation has been reported in the polymerase of yellow fever virus (a per genome rate of 0.002 [11], although this excludes lethal mutations, which, if included, could potentially increase the rate by a factor of two [12]). Furthermore, retroviruses have a broader range of rates than other RNA viruses (extending down to 0.06 mutations per genome per round of replication [13]). Perhaps more significantly, higher-fidelity RNA viral polymerases (which lead to a lower mutation rate) can be created *in vitro*. For example, repeated passage of poliovirus in the presence of the chemical mutagen ribavirin produced a mutant polymerase, differing by only one base, which showed higher replication fidelity than the wild type [14]. A ribavirin-resistant polymerase has also been selected for in FMDV [15]. Thus, RNA viruses can acquire a lower mutation rate, and it seems reasonable to assume that the option of reducing the mutation rate is open to natural populations, but is selected against.

There is also wide variation in the level of recombination, from several crossovers per round of replication in HIV [16] to effectively clonal replication in some negative-sense viruses [17]. This variation seems inconsistent with a struggle against harmfully high mutation rates, given that, in theory, recombination could alleviate a high mutational load [18].

Trade-off with replication speed

We believe that a more probable explanation for the high mutation rate of RNA viruses lies in a putative fitness cost to replication fidelity. Such a cost could be a reduced replication rate: viruses might be able to replicate either quickly or accurately, but not both [19].

Currently, there are few data on the relationship between these variables for RNA viruses. In support of this hypothesis, *in vitro* studies of the reverse transcriptase (RT) of HIV-1 showed a negative relationship between the rate of polymerisation and the rate of mutation [20], and vesicular stomatitis virus (VSV) clones with reduced mutation rates had a reduced competitive fitness (lower growth in cell culture compared with the wild type) [21]. There is also evidence for such a trade-off within DNA viruses: mutants of T4 phage differing by only one base pair in their DNA polymerase exhibited variation in mutation rate over four orders of magnitude [22]. This study found evidence for a cost to fidelity in the form of a reduced viral replication rate, with increased proofreading appearing to also lead to the removal of correct nucleotides.

Contrary to this hypothesis, another study of HIV-1 RT found that replacement of a methionine by a valine at one specific position in the same enzyme reduced the mutation rate, whereas replacement by an alanine increased it; however, both mutants showed a higher rate of polymerisation [23]. Also, a mutant poliovirus replicase with increased fidelity did not appear to have a reduced replication rate [24].

Consequences of a high mutation rate

We propose that the high mutation rate of RNA viruses has three main consequences for their evolution.

Population viability

Given that most mutations are harmful, high mutation rates might pose a problem for RNA viruses, and artificially raising the mutation rate even further could represent a viable antiviral strategy. This is the reasoning behind so-called lethal mutagenesis therapy [25,26]. The chemical ribavirin is used to treat several viral infections in humans including hepatitis C virus (HCV) and respiratory syncytial virus (RSV), and it is thought (although not conclusively demonstrated) that its effect is due to its being a known mutagen [27]. Chemical mutagens have also been shown to reduce the growth of at least another six RNA viral species in cell culture [25].

The idea that artificially elevating mutation rates could be a useful therapy is given weight by the existence of