Hepatitis C quasispecies adaptation in the setting of a variable fidelity polymerase

Daniel Schmidt-Martin¹
Orla Crosbie²
Elizabeth Kenny-Walsh²
Liam J Fanning¹

¹Molecular Virology Diagnostic and Research Laboratory, Department of Medicine, University College Cork, Clinical Sciences Building, Cork University Hospital; ²Department of Gastroenterology and Hepatology, Cork, Ireland

Abstract: Hepatitis C (HCV) is a virus characterized by an RNA-dependent RNA polymerase that lacks a proofreading mechanism and, as a result, generates a quasispecies. There is emerging evidence that this RNA-dependent RNA polymerase may in fact have variable fidelity. Here, we review the relevant concepts, including fitness landscapes, clonal interference, robustness, selection, adaptation, mutation rates, and their optimization, and provide a unique interpretation of a number of relevant theoretical models, evolving the theory of replicative homeostasis in light of their findings. We suggest that a variable fidelity polymerase can find its own optimal mutation rate, which is governed by the sequence itself and certain population dynamics. We propose that this concept can explain features of viral kinetics and clearance, both spontaneously and following treatment of chronic HCV. We point to evidence that supports this theory and explain how it refines replicative homeostasis and conclude by discussing particular areas of potential research that might augment our understanding of viral host interactions at an individual cellular level.

Keywords: fitness landscapes, adaptation, evolution, quasispecies, hepatitis C, replicative homeostasis

Introduction

Hepatitis C (HCV), a positive 9.2–9.6 kb RNA Flavivirus, was first identified by Choo and colleagues at Chiron in 1989 and is estimated to infect up to 3% people worldwide, equivalent to 120–170 million people.¹–³ Chronic HCV infection leads to the development of cirrhosis in 20% of cases after 20 years and is now the leading indication for orthoptic liver transplantation in the USA.

Low fidelity and the lack of proofreading ability of the HCV RNA-dependent RNA polymerase (RDRP) results in a population of closely related genomes or quasispecies.⁴ Originally proposed by Eigen as a model for the study of the evolution of primitive organisms, the quasispecies concept has been applied to many bacteria and viruses including human immunodeficiency virus and HCV.⁵ The gradual generation of point mutations results in the development of new variant species or “quasispecies” with slightly altered characteristics that then undergo selection. Within a given host, these quasispecies best adapted to the environment are most likely to survive and become dominant as a result of the principle of competitive exclusion.⁶ The most prevalent quasispecies is the “master” sequence and other related quasispecies cluster around this in terms of their genetic distance. Mutations either undergo selection (positive, resulting in the selection of beneficial traits; negative, when a deleterious trait is removed). Alternatively, in the absence of selection, the gradual accumulation of...
neutral or near neutral mutations of insufficient magnitude to prompt selection results in an evolutionary process known as “genetic drift.” Gradual adaptation to the host occurs as a result of these processes with the neutral theory of evolution predicting that genetic drift will be the predominant form of evolution. For a mutation to provoke a change from genetic drift to natural selection it must breach what has been termed the “selection threshold.”

These processes are dependent on a number of factors including virion fitness, population size, clonal density, clonal interference, and mutation rates. Our understanding of HCV has grown exponentially as a result of both experimental results and mathematical modeling, which have facilitated better understanding of viral replication processes and, as a result, viral genomic selection, adaptation, and evolution. This review provides an up-to-date appraisal of these topics.

Fitness
Conceptually, each genome has an inherent fitness defined by a group of characteristics (ability to infect, ability to replicate, energy requirement), with each quasispecies vying for host resources (host cells, cellular machinery, etc). Within the population, each sequence competes for these host resources with the best adapted, or fittest characteristics, most likely to dominate. However, the transience of this domination is guaranteed by the mechanism by which it is generated; the almost inexorable emergence of fitter mutants demands continual evolution for survival, in a process called the “Red Queen Hypothesis.” A moderate increase in viral fitness of one quasispecies over another results in exponential proliferation of this new quasispecies, with likely extinction of its competitor sequences in what amounts to a zero-sum game.

Fitness landscapes
Viral fitness can be described in the form of a fitness landscape, with mountains corresponding to areas of increased fitness surrounded by areas of diminishing fitness analogous to foothills (Figure 1). The accumulation of mutations allows the exploration of the sequence space and through this process the discovery of fitness gains that might displace the master sequence through competition. In the case of HCV, because the number of possible nucleotide combinations is so great ($4^{600}$), this landscape is only able to describe quasispecies diversity for short segments of the sequence. Not all mutations are viable, with these lethal mutations akin to cliffs in what is known as a “truncated fitness landscape.” Finally, the combined interplay between individual quasispecies and the immune system results in a changing or dynamic truncated fitness landscape.

In this setting, the lack of a proofreading function is often looked upon as beneficial to HCV; adapted mutants, which are closely related to the parent virion and better able to evade the host’s immune response, emerge and maintain chronic infection. However, in this model the proviso is that high mutation rates mean that beneficially adapted mutants are equally prone to deleterious mutations, which can potentially wipe out entire quasispecies. Muller’s ratchet predicts that deleterious mutations are likely to “hitchhike” and be found in all future progeny, barring the unlikely event of a reciprocal mutation taking place. Mitigating the effects of hitchhiking is the process of recombination, which can facilitate the removal of deleterious mutations by combining mutation-free segments and allow greater potential exploration of the sequence space by combining sequences with multiple mutations. It is this latter process that is thought to contribute significantly to the emergence of differing HCV genotypes and even taxa and species.

Interference
Although a significant factor in determining the fate of a given quasispecies, competitive exclusion is not the sole determinant of evolutionary success. In large quasispecies populations, it has been shown that sequences with significant fitness superiority are not necessarily guaranteed to dominate a quasispecies due to a process known as “clonal interference.” In small populations, beneficial mutations of smaller increments are more likely to come to dominate as a result of selective sweep, while, in large populations (as are seen in established chronic HCV infections), a quasispecies with significant fitness benefit can be suppressed by the less-fit dominant quasispecies, unless it reaches a

![Figure 1 Schematic representation of the sequence space in the form of a fitness landscape.](image-url)
critical threshold. This has the net effect of ensuring that, in chronic infection, the incremental increase in quasispecies fitness becomes larger in fitness gain but more infrequent in occurrence. Experimental evidence of clonal interference supporting this theory has been found in *Escherichia coli*, DNA viruses, HCV, and the RNA vesicular stomatitis virus (VSV).

**Defective interfering particles**

Notwithstanding the extreme variability seen in the genetic sequence of RNA viruses, it must be remembered, however, that redundancy in the sequence is limited and that the proteins produced are small in number and, in most cases, essential in function. However, despite this lack of redundancy, subgenomic particles exist that can have significant effects on virus population dynamics.

Named “defective interfering particles” (DIPs) and identified in several virus species (including both DNA and RNA viruses), they may be important factors in the search for fitter quasispecies resistant to the effects of DIPs. Unable to replicate in the absence of wild-type virus but able to infect new cells, they are thought to contribute to the oscillating nature of the viral load repeatedly seen in HCV infection. DIPs have also been proposed to interfere in the production of wild-type virus and modulate pathogen virulence and may themselves be potential antiviral agents.

Stumpf and Zitzmann have proposed the reciprocality of DIPs; that is, that the particles are able to replicate but are unable to cause de novo infection of new cells due to the deletion of the structural section of the genome. The associated increase in replicative ability leads to competitive exclusion of viable virions and the gradual accumulation of defective intracellular viral RNA, meaning that continuous de novo infection of new cells is essential to viral survival. Experimental evidence for this has remained elusive.

Studies focused on hepatocyte-derived HCV genomic sequences have not found evidence of these particles, though factors such as the duration of infection and use of limited numbers of clones (it is estimated that use of 20 clones will demonstrate most sequences present at a level of 10%) may go some way to explain this. Indeed, the advent of next-generation sequencing may see the reemergence of this concept.

**Robustness**

The ability of a virion to tolerate mutations without phenotypic disruption, termed “robustness,” is also likely to be important in maintaining or enhancing fitness. Characterized by a greater number of available neutral mutations, a high degree of robustness results in a smoother fitness landscape, in a theory described as “survival of the flattest.” Studies using digital models and subviral particles suggest that an organism with greater robustness may out compete and dominate less robust counterparts, particularly at times of high mutagenesis.

The emergence of neutral mutation-rich organisms may however have significant implications for virion evolution. A recent study has demonstrated that a high proportion of neutral or near-neutral mutations may act as a barrier to evolution by natural selection, with genetic drift coming to dominate. Studies evaluating HCV robustness are limited, but one network-based analysis of HCV polyprotein has demonstrated a high degree of robustness at many nucleotide positions, with relatively few positions vulnerable to phenotypically deleterious mutation. Comparisons with other RNA viruses are challenging, as direct studies have not yet been published. One recent paper has estimated by site-directed mutagenesis that 40% of random mutations in VSV are lethal, which may suggest a lesser degree of robustness compared with HCV.

Finally, it has been suggested that increased robustness may result in a reduced ability to adapt and that, in organisms that are required to survive in changing environments, the requirement for frequent adaptive change will limit tolerance of neutrality/robustness. Indeed, the ability of an organism to respond to selective pressure and tolerate significant large-scale genetic evolution or evolvability has also been demonstrated as a selectable trait.

**Cooperative interaction**

The concept of the “cooperative interaction” of the constituent mutants in exploring fitness maxima, so that the population ultimately achieves a mutation–selection equilibrium, distinguishes quasispecies theory from classical population genetics. When looked on in this light, it becomes apparent that successful quasispecies evolution is a population-wide phenomenon, so that fitness can be seen as an “ensemble property.” While evidence for this phenomenon is limited, studies of poliovirus have demonstrated that the pathogenesis of individual quasispecies is affected by cooperative interaction with other mutants in the quasispecies profile and that maintenance of a degree of heterogeneity is preferable for viral survival and maintenance of tissue tropism. Indeed, the influence of cooperative reactions has been proposed as essential if mathematical models are to accurately generate the quasispecies patterns observed in vivo.
Adaptation

“Adaptation” is the process whereby the quasispecies alters to become more suited to new or changing environments. The rate of adaptation of quasispecies appears to be governed by a number of factors: population size, mutation rate, adaptive quotient, and variability of the environment. In small populations, the size of the population limits the ability to explore the sequence space. As a result, adaptation occurs at a slower rate by means of stochastic genetic drift with episodic selective sweeps. This means that the population is more likely to be confined to local fitness peaks. In contrast, large populations are better able to expand throughout the sequence space and, as a result, adaptation is more deterministic, though the time taken for fixation of beneficial mutations is increased as a result of interference.45,46

Mutation rates

The effects of different mutation rates on a quasispecies within a truncated fitness landscape appear to follow three patterns: (1) low mutation rates result in a distribution around the master sequence and are more likely to become “trapped” in local fitness peaks, reducing the chances of complete exploration of the sequence space; (2) intermediate mutation rates result in wider exploration of the sequence space, with the emergence of variants further removed from the master sequence; and (3) those with high mutation rates produce an ever-increasing number of progeny with lethal mutations and, as a result, reach what has been called the “error threshold” – the point at which the quasispecies becomes unable to maintain sequence integrity. The coercion of viruses beyond this point into what has been called “error catastrophe” has been a major strategy in the development of antiviral therapies.46–49

Interestingly, adaptation is not maximal prior to reaching error threshold; rather, it behaves in a sine wave fashion (Figure 2). The mutation rate that results in optimal adaptation is remote from error threshold and adaptation decreases with increasing mutation rate as the ability to fix beneficial mutations decreases until error threshold is breached.50,51 Using parameters present in quasispecies, Orr has found that optimal mutation rates for adaptation are governed by the strength of selection against deleterious mutations (ie, more truncated landscapes have a lower optimal adaptive mutation rate).50 This work was based on the assumption that the selective power against deleterious mutations was at all times greater that the selective power for beneficial mutations; however, this may not, in fact, be the case.

Adaptive quotient

Johnson and Barton advanced this theory by describing a matrix that can predict a sequence-specific optimal mutation rate depending on whether the surrounding fitness landscape is dominated by deleterious or beneficial mutations and the selective power of these relative to each other.52 According to this model, the existence of many potential beneficial mutations will promote the emergence of a higher mutation rate and vice versa. In the setting of a bottleneck event (rapid reduction in quasispecies – as occurs at transmission of HCV) the organism can be seen to be less adapted to the new host and, as a result, the ratio of beneficial:deleterious mutations is also likely to change and will probably be reflected in the rate of mutation.53

Variability of environment

In static environments, the exploration of the sequence space with fixation of beneficial mutations that pass the selection threshold and outlast clonal interference becomes exhausted once the quasispecies reaches the mutation–selection equilibrium. At this stage, all fitness optima have been explored. As this occurs, the fitness gains that were initially large, diminish toward nil.54 With many microorganisms, however, the emergence of new environments, either as a result of transmission of infection or the development of immune responses, results in a dynamic fitness landscape that serves to replenish the potential for adaptive change.

Furthermore, the ruggedness of these landscapes can themselves affect the rate of adaptation. Clune et al demonstrated using computer models that digital organisms fail to optimize mutation rates and tend to settle at a mutation rate below this.55 Evidence for this has been described in DNA bacteriophages, where the imposition of a fourfold increase in mutation rate actually conferred fitness gain.55,56 Clune et al argued that, while adaptation occurs over long periods (many generations), selection acts quickly and this phenomenon
may be an effort by the virus to mitigate the potential for emergence of deleterious/lethal mutations. Expanding on this initial finding, Clune et al demonstrated that the observed mutation rate is dependent on the ruggedness of the fitness landscape with more rugged landscapes favoring an even lower mutation rate.55

**The fidelity spectrum**

There is a growing body of literature indicating that mutation rates are not constant and may be selectable. Mutation rates have been shown to increase at times of stress in many bacteria and lethal mutagenesis has long been suggested as a potential treatment strategy in viral infections.57 The beneficial effects of increased mutational rates, in addition to how they may be associated with increased replicative capacity, have also been demonstrated in bacteriophage populations.58 Furthermore, adaptive change in the mutation rate in response to medications has been shown to confer drug resistance and sustain chronic infection in the case of human immunodeficiency virus type 1.59 Several mechanisms governing how transient increases in mutation rate can be generated and suitably regulated have been suggested, including environmental and heritable factors.60

**Evidence of variable RDRP mutation rates in HCV**

In HCV, the estimated mutation rate is $1 \times 10^{-4}$ to $5 \times 10^{-5}$/base.61-63 The estimation of the error threshold of HCV RDRP is $10^{2}$ to $10^{3}$ (mutations per base), which leaves scope for a ten- to hundredfold change in baseline RDRP fidelity before the error threshold is reached, with the optimal adaptive mutation rate likely to be found within this range. Lethal mutagenesis has formed one of the theories of the mechanism of action of ribavirin, as it has been shown to induce lethal mutagenesis in poliovirus and foot-and-mouth disease, but the results in HCV have been variable.62,64,65 Ribavirin-resistant mutations have been described, both in vitro and in vivo, in two HCV nonstructural proteins (NS5A and NS5B) including the RDRP. It has been suggested that the NS5A mutation may indicate that this protein may interact with the RDRP to modulate polymerase fidelity.62,66,67 Indeed, the idea that RDRP fidelity may be controlled remotely is not novel to HCV.68

Recently, mutant RDRPs conferring ribavirin resistance by means of increased fidelity have been described in both foot-and-mouth disease virus and poliovirus. We feel that similar mutants are likely to exist in the case of HCV and that their emergence during ribavirin therapy would explain the heterogeneity of the effect on mutation rates seen in these studies. Furthermore, the sampling intervals may have been such that transient increases in mutation rates may have been missed.

A mechanism by which viruses might self-regulate replication fidelity has been proposed by Sallie in his theory of replicative homeostasis (RH).69-71 Sallie argued that HCV viral kinetics behave in such a way to suggest autoregulation of virion production through a homeostatic mechanism that modulates RDRP fidelity/processivity (which he proposed are inversely proportional). RH predicts that excess wild-type protein will prompt a decrease in fidelity and a resultant increase in mutation and vice versa. The idea that mutation rates may be dependent on polymerization rates was first proposed in the kinetic proofreading hypothesis, in which a delay in the rate of polymerization results in increased polymerase fidelity; experimental support for this has been demonstrated in the case of VSV.11,72 While we are of the opinion that Sallie’s theory has significant merit, we feel that the theory of RH could be further adapted to more accurately describe the behavior of HCV.

**A framework for the action of a variable fidelity polymerase**

On the basis of the position of the quasispecies within the fitness landscape, we propose that a framework exists for the selection of phase-specific mutation rates (Figure 3). The RDRP acts along a fidelity spectrum with optimal mutation rates that are largely dependent on population size, capacity for adaption (adaptive quotient), and variability of
We propose that the optimal mutation rate selected for could be predicted by the position of the sequence within a framework similar to that proposed by Johnson and Barton. The exploration of this fidelity spectrum is likely to be initially stochastic, as it is reliant on the generation of promutator mutations and evidence for similar processes can be seen in Drosophila populations.

Following the bottleneck of transmission, unencumbered by clonal interference, and with an increased probability of beneficial mutations, we suggest that a form of density-dependent selection, similar to those that have been described in foot-and-mouth disease virus, E. coli, and Drosophila, will result in the emergence of an increased mutation rate.

This latter occurrence would be characterized by quicker adaptation, could correspond to the intermediate fidelity phase as described by Saakian et al. and, could be likened to the episodes of stress which have also been shown to result in increased mutation rates in E. coli. As the relative proportion of beneficial to deleterious mutations is increased in small nonadapted populations, the emergence of an increased mutation rate is favored. Initial infection with a finite number of variants will gradually explore local fitness maxima by stochastic means until the population becomes sufficient for deterministic exploration as the capacity to generate all possible mutants is achieved. With population expansion, increasing clonal interference, and viral adaptation, the same process will select a less-productive polymerase with increased fidelity that has the added potential bonus of being immunologically stealthy by means of viral-load reduction. This period of selection may result in the reduction in viral load often seen in acute HCV.

Furthermore, increased fidelity will inevitably mean that antigenic thresholds will be intermittently breached, resulting in activation of the adaptive immune response. As the exploration of both sequence space and what we refer to as the “fidelity spectrum” is stochastic, it is to some degree dependent on chance, but the near certainty of successful exploration has been built into the quasispecies characteristics of the virus. Conversely, the certainty of failure, in some cases, to either optimize fidelity or even find infidelity sufficient to evade the immune response, provides the tantalizing prospect of explaining the process by which infection is cleared in 15%–25% of patients. The emergence of a population selection–mutation equilibrium will tend toward a lower mutation rate, as the genetic distance to the nearest beneficial mutations is likely to become larger due to this adaptation. In summary, at times of stress, the polymerase and its inherent mutation rate becomes the unit of selection, while, at other times, it is the genomic properties and their cooperative/competitive interactions that become the traits selected for or against.

In our model, similar to that of RH, the selection of particular sequences for removal by the immune system will merely result in the generation of new quasispecies to match the new fitness landscape, while also facilitating long-term stability of quasispecies in the absence of variations in effective immune pressure. This model also has the capacity to explain HCV clearance in the absence of seroconversion, as it allows for the attainment of error catastrophe without the need for immune response. Additionally, our proposed mechanism of action along a fidelity spectrum more coherently explains why the emergence of a single dominant quasispecies in the treatment of HCV infection and a low rate of quasispecies evolution are more likely to result in clearance as opposed to the generation of new quasispecies, as Sallie’s model would suggest.

In proposing this model, we must acknowledge that one of the major obstacles to clarifying the interaction between quasispecies theory and experimental results in HCV is the phenomenon of founder effect at the level of the individual cell. The prevention of superinfection, in theory, means that the apparatus of the cell is at the mercy of this sole founder and that competition is prevented, promoting the preservation of the status quo. This, coupled with evidence demonstrating the prevention of infection of neighboring cells via the apical cell membrane and the facilitation of virion transfer to these neighbors via tight junctions, is equivalent to dynasty building—that is, clonal expansion. Accounting for these factors in evolutionary models is challenging, particularly when little is known of the incidence of superinfection in the context of fitter “pilgrim” virions, which may facilitate the conversion of the quasispecies to new fitness optima.

Finally, we would like to note one conundrum reconciling the current theories of optimal mutation rates and the suggestion that organisms adapt toward neutral networks. Under the survival of the flattest hypothesis, the emergence of such fitness landscapes results in a reduction in the ruggedness of the fitness landscape. As the number and selective power of potentially deleterious mutations are reduced, we should see closer optimization of mutation rates to maximize adaptation. However, little evidence has been produced to favor this and, conversely, the mutation rate in E. Coli, which has a 90% tolerance of mutations, has a mutation rate far less than that of RNA viruses, which have a lethal mutation rate of 21%–40%.
Conclusion
Adaptive evolution is slave to both genetic drift and natural selection, with the emergence of more neutral flatter fitness landscapes favoring the former. Following a bottleneck, the exploration of the sequence space is stochastic, with the transition to deterministic exploration dependent on the population size and the development of clonal interference. Mutation rates often fail to optimize adaptation and this may be an effort to mitigate the relative strength of lethal mutations when compared with the relative and often-marginal benefit of beneficial mutations – particularly, in well-already-adapted species. Mutation rates are not constant and, in low population sizes, increased mutation rates may be selected for to enhance the rate of adaptation. Several potential mechanisms for regulating mutation rates to ensure that these increases are transient have been proposed. HCV demonstrates characteristics consistent with a population density-mediated selection of mutation rates.

Acknowledgment
The authors would like to acknowledge the support of Daniel Schmidt-Martin from Molecular Medicine Ireland.

Disclosure
The authors report no conflicts of interest in this work.

References


