Published in final edited form as:

Bull Math Biol. 2012 August ; 74(8): 1789–1817. doi:10.1007/s11538-012-9736-y.

Modeling quasispecies and drug resistance in hepatitis C patients treated with a protease inhibitor

Libin Rong^{†,‡}, Ruy M. Ribeiro[‡], and Alan S. Perelson^{‡,*}

[†]Department of Mathematics and Statistics and Center for Biomedical Research, Oakland University, Rochester, MI 48309

[‡]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545

Abstract

Telaprevir, a novel hepatitis C virus (HCV) NS3-4A serine protease inhibitor, has demonstrated substantial antiviral activity in patients infected with HCV. However, drug-resistant HCV variants were detected in vivo at relatively high frequency a few days after drug administration. Here we use a two-strain mathematical model to explain the rapid emergence of drug resistance in HCV patients treated with telaprevir monotherapy. We examine the effects of backward mutation and liver cell proliferation on the preexistence of the mutant virus and the competition between wild-type and drug-resistant virus during therapy. We also extend the two-strain model to a general model with multiple viral strains. Mutations during therapy only have a minor effect on the dynamics of various viral strains, although they are capable of generating low levels of HCV variants that would otherwise be completely suppressed because of fitness disadvantages. Liver cell proliferation may not affect the pretreatment frequency of mutant variants, but is able to influence the quasispecies dynamics during therapy. It is the relative fitness of each mutant strain compared with wild-type that determines which strain(s) will dominate the virus population. This study provides a theoretical framework for exploring the prevalence of preexisting mutant variants and the evolution of drug resistance during treatment with other HCV protease inhibitors or polymerase inhibitors.

Keywords

Telaprevir; mutation; fitness; quasispecies; direct-acting antiviral agents; mathematical model

1 Introduction

Chronic infection with hepatitis C virus (HCV) has caused an epidemic with approximately 130 to 170 million people infected worldwide and 3 to 4 million individuals newly infected each year [1]. About 80% of newly infected patients develop chronic infection. Of those chronically infected, 60 – 70% develop chronic liver disease, 5 – 20% develop cirrhosis, and 1 – 5% die from cirrhosis or liver cancer [1]. A combination of pegylated interferon (PEG-IFN), administered once weekly, and daily oral ribavirin (RBV) has been used to treat HCV infection for 24 or 48 weeks [2]. Although the combination exerts synergistic antiviral effects [3, 4], it leads to sustained viral elimination in only some treated patients. The HCV genotype appears to be an important factor in predicting response. In patients infected with genotypes 2 and 3, about 80–90% of patients achieve a sustained viral response (SVR), defined as the absence of detectable serum HCV RNA 24 weeks after completion of treatment. In patients infected with genotype 1, the major genotype affecting North America,

*Corresponding author: asp@lanl.gov.

Europe and Japan, only about 40% of treated individuals achieve SVR [2]. Lack of a complete response, viral relapse following treatment, and premature termination of therapy due to adverse events that occur during dosing all contribute to this unsatisfactory response rate observed among HCV genotype 1 infected patients. Therefore, new antiviral drugs with higher efficacy, shorter treatment duration, and a more favorable side-effect profile as a monotherapy or in combination with other antivirals are highly desirable.

New treatment options are focused on the development of direct-acting antiviral agents (DAAs) that target different steps of the HCV life cycle [5–7]. An important target is the HCV-encoded NS3-4A serine protease. In clinical trials HCV protease inhibitors have been used to treat HCV genotype 1 infected patients. They have shown an impressive capacity to block the NS3-4A protease-dependent cleavage of the HCV polyprotein, which is an essential step in viral replication (HCV replication will be discussed in detail later). The first protease inhibitor, BILN 2061 (ciluprevir; Boehringer-Ingelheim), showed potent antiviral activity in patients infected with HCV genotype 1 [8], but clinical development was halted due to drug-induced cardiotoxicity [9]. Boceprevir (SCH 503034; Schering-Plough Pharmaceuticals), another oral HCV protease inhibitor, also demonstrated substantial antiviral effects when used in combination with PEG-IFN- α -2b in HCV genotype 1 infected patients, who were previously nonresponders to PEG-IFN- α -2b with/without RBV therapy [10]. Telaprevir (VX-950; Vertex Pharmaceuticals) is a reversible, selective, and specific peptidomimetic inhibitor of NS3-4A that is effective in inhibiting viral replication in HCV replicon cells [11]. It had a favorable pharmacokinetic profile with high exposure in the liver in several animal models [12], and in monotherapy induced a profound decline of plasma HCV RNA levels of the order of 3–4 logs in patients infected with HCV genotype 1 treated for 14 days [13]. Both boceprevir and telaprevir have recently been approved by the US Food and Drug Administration (FDA) to treat HCV infection when used in combination with PEG-IFN and RBV. The addition of either of them to therapy with PEG-IFN and RBV has significantly increased the rates of SVR [14–18].

Emergence of drug-resistant mutations is a problem challenging the development of direct-acting antiviral drugs. Like most RNA viruses, HCV evolves rapidly because of high-level viral replication through an error-prone RNA polymerase that lacks associated proofreading capacity. As a consequence, the viral population exists as a complex mixture of genetically distinct, but closely related, variants commonly referred to as a quasispecies, whose composition is subject to continuous change due to the competition between newly generated mutants and existing variants with different phenotypes and fitness [19]. During antiviral therapy pre-existing minor viral populations with reduced susceptibility to the administered drug or drugs will gain a growth advantage over wild-type and become the dominant genotype. The amino acid substitutions selected by protease inhibitors that confer drug resistance have been characterized *in vitro* in the HCV replicon system [20–22].

The initial selection and kinetics of telaprevir-resistant HCV variants have been further described in patients given the protease inhibitor alone [13, 23] or in combination with PEG-IFN- α -2a [24, 25]. Although 14 days of treatment resulted in substantial decreases in HCV RNA levels, there was evidence of viral breakthrough in some patients during the dosing period, which was believed to be associated with the selection of HCV variants with reduced sensitivity to telaprevir [13]. Using a highly sensitive sequencing assay, Sarrazin et al. [23] identified mutations that confer resistance to telaprevir in the NS3 protease catalytic domain and correlated them with virologic response. These mutations were further investigated in a subsequent study [25] that provides a detailed kinetic analysis of HCV variants in patients treated with telaprevir alone or in combination with PEG-IFN- α -2a for 14 days. The four HCV genotype 1a infected patients in the telaprevir monotherapy group all exhibited viral load rebound during the dosing period. Virus isolated from these

patients at day 2 contained low levels (5%–20%) of single-mutant resistant variants, which increased in the population of virus isolated at days 6 and 10, and were replaced by more resistant double-mutant variants by day 13 and during the first follow-up week with PEG-IFN plus RBV [25]. Why drug-resistant viral variants emerged so rapidly following treatment with telaprevir is not fully understood.

In this paper, we study HCV quasispecies and drug resistance in patients treated with the protease inhibitor telaprevir. We begin with a simple two-strain model in which liver cells, e.g., hepatocytes, infected with wild-type virus are able to produce not only wild-type virus but also a small amount of drug-resistant variants. The two-strain model was studied numerically and was shown to fit the observed dynamics of both drug-sensitive and drug-resistant viruses in patients treated with telaprevir [26]. Here we study this model analytically. With reasonable simplifications, we obtain an analytical solution for the mutant frequency in patients given telaprevir alone, which is capable of explaining the rapid selection of pre-existing drug resistant variants after therapy initiation. We study the competition between wild-type and drug-resistant virus during treatment. We also examine the effects of backward mutation and hepatocyte proliferation on the pre-existing mutant frequency and the evolution of viral variants during therapy. Extending the two-strain model, we then develop a multi-strain model in which drug-resistant HCV variants that differ at more than one site are incorporated. We calculate the expected frequency of each viral strain in untreated patients. The results of the competition between multiple viral variants during therapy with telaprevir are also provided. Because telaprevir and boceprevir inhibit the same HCV protease, the analysis in this study with telaprevir can be applied to boceprevir or to other HCV protease inhibitors under development.

2 Rapid emergence of drug resistance

2.1 Model description

Before describing the model, we use a diagram of the HCV life cycle (Figure 1) as a framework for discussing our current knowledge of virus replication. The exact mechanism by which HCV enters hepatocytes, the primary targets of infection, is still largely unknown. It is presumably receptor-mediated and involves CD81 [27], the human scavenger receptor class B type 1 (SR-B1) [28], and other molecules such as claudin-1 [29] and occludin [30]. Following fusion of the viral and cellular membranes, nucleocapsid enters the cytoplasm of the host cell and releases a single-stranded, positive-sense RNA genome (uncoating). This genome serves, together with newly synthesized RNAs, multiple roles within the HCV life cycle: as a messenger RNA (mRNA) for translation to produce a large polyprotein, as a template for HCV RNA replication, and as a nascent genome that is packaged in progeny virus particles. The generated polyprotein is then cleaved by several enzymes including the NS3-4A serine protease to produce 10 viral proteins: the structural proteins (the core protein C, glycoproteins E1 and E2), a small integral membrane protein p7, and the nonstructural (NS) proteins NS2, NS3, NS4A, NS4B, NS5A and NS5B. This is followed by RNA replication that occurs in a specific cytoplasmic membrane alteration, termed the “membranous web”, whose formation is induced by the integral membrane protein NS4B [31]. The process of RNA synthesis is not fully characterized, but is likely to be semi-conservative and asymmetric [32]: the positive-strand genome RNA serves as a template for the synthesis of a negative-strand intermediate; the negative-strand RNA then serves as a template to produce multiple nascent genomes. Both of these steps are catalyzed by the NS5B RNA-dependent RNA polymerase (RdRp). In the meantime, structural proteins E1, E2 and C have matured. Together with progeny positive-strand genomes, they assemble and are ready for vesicle fusion at the host cell plasma membrane, after which new virions are released into the extracellular milieu by exocytosis.

The viral RdRp is an important enzyme that catalyzes the synthesis of both positive- and negative-strand RNAs. However, the HCV RdRp has a high error rate, with a misincorporation rate of 10^{-4} – 10^{-5} per copied nucleotide [33]. Furthermore, since the RdRp is devoid of proofreading capacity and other postreplicative repair mechanisms, it cannot correct misincorporations that occur randomly during replication [19]. The high mutation rate, together with rapid HCV replication [34] and the large viral population size, results in progressive diversification of viral genotypes and subtypes in geographically or epidemiologically-linked populations, and in the quasispecies nature of the virus population in a given infected individual [19].

We adapt a mathematical model, which was used to study HIV-1 infection and drug resistance [35], to examine the quasispecies dynamics of HCV before and during treatment. Based on the error-prone nature of the HCV polymerase, hepatocytes infected with wild-type virus are expected to produce both wild-type virus and mutant variants. A simple model including two strains, wild-type and drug-resistant (assuming a single mutation confers a certain level of drug resistance), is described by the following equations:

$$\begin{aligned}\frac{d}{dt}T(t) &= s - dT - \beta_s V_s T - \beta_r V_r T, \\ \frac{d}{dt}I_s(t) &= \beta_s V_s T - \delta I_s, \\ \frac{d}{dt}I_r(t) &= \beta_r V_r T - \delta I_r, \\ \frac{d}{dt}V_s(t) &= (1-\mu)p_s I_s - cV_s, \\ \frac{d}{dt}V_r(t) &= \mu p_s I_s + p_r I_r - cV_r,\end{aligned}\tag{1}$$

where T is the number of target cells; I_s and I_r are the numbers of cells infected with wild-type and drug-resistant virus, respectively; V_s and V_r represent the numbers of wild-type and drug-resistant virus, respectively. Target cells are produced at rate s and die at rate d . Cells become infected with wild-type virus at rate β_s , and infected with drug-resistant virus at rate β_r . Once infected, cells die with death rate δ . HCV virions are produced at different rates, p_s and p_r , by infected cells, I_s and I_r , respectively, while the two strains have the same virion clearance rate c . Taking a single mutation into account, we assume that I_s with a probability μ produces drug-resistant virus. μ is about 10^{-4} – 10^{-5} per copied nucleotide. We note that model (1) is different from the two-strain model developed during HIV treatment [35] because mutation in HCV occurs during the production of virus rather than at infection like HIV-1. Thus, an infected cell can produce a spectrum of viral variants (a model with multiple viral strains is discussed later). Backward mutation from mutant to wild-type is neglected here, but will be incorporated into the model for comparison later.

2.2 The frequency of the mutant virus before treatment

There are 3 possible steady states $(\bar{T}, \bar{I}_s, \bar{I}_r, \bar{V}_s, \bar{V}_r)$ of model (1): the infection-free steady state $E_0 = (s/d, 0, 0, 0, 0)$, the boundary steady state $E_r = (\frac{c\delta}{\beta_r p_r}, 0, 0, (\mathcal{R}_r - 1)\frac{dc}{\beta_r p_r}, (\mathcal{R}_r - 1)\frac{d}{\beta_r})$, and the coexistence steady state $E_c = (\bar{T}, \frac{\beta_s}{\delta}\bar{T}\bar{V}_s, \bar{V}_s, \frac{\beta_r}{\delta}\bar{T}\bar{V}_r, \bar{V}_r)$, where

$$\bar{T} = \frac{c\delta}{(1-\mu)\beta_s p_s}, \quad \bar{V}_s = \frac{[(1-\mu)\mathcal{R}_s - 1]d}{\beta_s + \beta_r \frac{\mu}{1-\mu-r}}, \quad \bar{V}_r = \frac{\mu}{1-\mu-r}\bar{V}_s, \quad r = \mathcal{R}_r / \mathcal{R}_s, \quad \mathcal{R}_s = s\beta_s p_s / (dc\delta) \text{ and } \mathcal{R}_r = s\beta_r p_r / (dc\delta)$$

are the basic reproductive ratios of the wild-type strain and the drug-resistant strain, respectively. The ratio r represents the relative fitness of drug-resistant to wild-type virus. The steady state E_r exists if and only if $\mathcal{R}_r > 1$ and E_c exists if and only if $\mathcal{R}_s > \max(1/(1-\mu), \mathcal{R}_r/(1-\mu))$.

The above existence conditions also provide threshold conditions for the stability of the steady states. Indeed, we can show that (i) when $\mathcal{R}_s < 1/(1-\mu)$ and $\mathcal{R}_r < 1$, E_0 is locally asymptotically stable; (ii) when $\mathcal{R}_r > 1$ and $r > 1-\mu$, E_r is locally stable; and (iii) when $\mathcal{R}_s >$

$1/(1 - \mu)$ and $r < 1 - \mu$, E_c is locally stable. The predominance of wild-type virus before treatment suggests $\mathcal{R}_s > 1$ and resistance-associated loss of fitness implies $\mathcal{R}_r < \mathcal{R}_s$, i.e., $r < 1$ [23]. These observations suggest that the conditions in (iii) are satisfied because μ is very small. As a consequence, the solutions of model (1) typically converge to the steady state E_c , i.e., both wild-type and drug-resistant viral strains coexist in infected individuals before therapy, as has been shown experimentally [36].

We calculate the frequency of the pre-existing drug-resistant variants in the total virus population from the coexistence steady state E_c . The mutant frequency is given by $\Phi = \tilde{V}_r / (\tilde{V}_s + \tilde{V}_r)$, where $\tilde{V}_s$ and $\tilde{V}_r$ are the steady states of wild-type and drug-resistant virus, respectively. Φ can be simplified to $\Phi = \mu / (1 - r)$, where $r = \mathcal{R}_r / \mathcal{R}_s$. Therefore, the mutant frequency before therapy depends only on the mutation rate and the relative fitness between mutant and wild-type virus. Since μ is small, the mutant variant remains at a very low level, although it coexists with wild-type in patients before treatment. We also note that the mutant frequency obtained above is equivalent to the result in population genetics where the mutant frequency is derived by mutation-selection balance [37], i.e., the frequency of a deleterious allele is approximately equal to the mutation rate (μ) divided by the selection coefficient (equivalent to $1 - r$ in the expression of Φ). If the mutant frequency is very low then stochastic models rather than this deterministic model would be needed. Here we only deal with mutants at frequencies that are high enough that deterministic models should be applicable.

2.3 Increase of the mutant frequency following treatment

The HCV NS3-4A serine protease plays an important role in viral polyprotein processing, cleaving at the NS3-4A junction and all downstream sites. Telaprevir, a new protease inhibitor, has been developed to block this step in the viral life cycle [12] and has been shown to profoundly reduce the plasma viral load in infected individuals [13, 23]. This is not surprising since the products of polyprotein cleavage are needed to mediate viral RNA replication and virion assembly (Figure 1). Assuming ε_s and ε_r are the drug efficacies of telaprevir in blocking viral production for wild-type and drug resistant virus, respectively, where $0 \leq \varepsilon_s, \varepsilon_r \leq 1$, the model under treatment with the protease inhibitor becomes

$$\begin{aligned} \frac{dT}{dt} &= s - dT - \beta_s V_s T - \beta_r V_r T, \\ \frac{dI_s}{dt} &= \beta_s V_s T - \delta I_s, \\ \frac{dI_r}{dt} &= \beta_r V_r T - \delta I_r, \\ \frac{dV_s}{dt} &= (1 - \mu)(1 - \varepsilon_s) p_s I_s - c V_s, \\ \frac{dV_r}{dt} &= \mu(1 - \varepsilon_s) p_s I_s + (1 - \varepsilon_r) p_r I_r - c V_r. \end{aligned} \quad (2)$$

Assuming T remains at the pretreatment steady state level, $T_0 = c\delta / [(1 - \mu)p_s\beta_s]$, over a short period of time after drug administration, and ignoring the term $\mu(1 - \varepsilon_s)p_s I_s$ in the V_r equation (because μ is $\sim 10^{-4}$ – 10^{-5} [33] and ε_s is close to 1 [38]), we can reduce Eq. (2) to a simpler system. The solution of the simplified system is $V_s(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$, $V_r(t) = C_3 e^{\lambda_3 t} + C_4 e^{\lambda_4 t}$, where λ_i , $i = 1, 2, 3, 4$, are the eigenvalues of the system, given by $\lambda_{1,2} = -[c + \delta \pm \sqrt{\Delta_1}] / 2$ and $\lambda_{3,4} = -[c + \delta \pm \sqrt{\Delta_2}] / 2$, with $\Delta_1 = (c + \delta)^2 - 4[c\delta - (1 - \varepsilon_s)(1 - \mu)p_s\beta_s T_0]$ and $\Delta_2 = (c + \delta)^2 - 4[c\delta - (1 - \varepsilon_r)p_r\beta_r T_0]$, which using the expression of T_0 can be simplified to $\Delta_1 = (c + \delta)^2 - 4\varepsilon_s c\delta$ and $\Delta_2 = (c + \delta)^2 - 4c\delta \left[1 - \frac{(1 - \varepsilon_r)\beta_r}{(1 - \mu)\beta_s} \right]$. In Appendix A, we show that $\lambda_i < 0$ and $C_i > 0$ for $i = 1, 2, 3, 4$.

The mutant frequency following treatment is then given by the following function of t ,

$$\Phi(t) = \frac{V_r(t)}{V_s(t) + V_r(t)} = \frac{C_3 e^{\lambda_3 t} + C_4 e^{\lambda_4 t}}{C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + C_3 e^{\lambda_3 t} + C_4 e^{\lambda_4 t}}, \quad (3)$$

which depends on c , δ , μ , ε_s , ε_r , r , and the time t since therapy began.

To study the change of the mutant frequency $\Phi(t)$ after drug administration, we have to determine the drug efficacy of telaprevir for each strain. The effectiveness of a drug against

wild-type virus can be approximated by a simple function [39] $\varepsilon_s(t) = \frac{C(t)^h}{IC_{50}^h + C(t)^h}$, where $C(t)$ is the drug concentration, IC_{50} is the concentration of drug needed to inhibit viral production by 50%, and h is the Hill coefficient. Based on the pharmacodynamics of telaprevir, the drug efficacy for the wild-type strain was calculated to be 0.9997 (median) [40], which is consistent with the 3–4 log first-phase drop of plasma HCV RNA levels when telaprevir was administered in monotherapy [13]. Similarly, for the mutant virus with n -fold

resistance, i.e., an n -fold increase in IC_{50} , we have $\varepsilon_r(t) = \frac{C(t)^h}{(n \cdot IC_{50})^h + C(t)^h}$. From the

equations of $\varepsilon_s(t)$ and $\varepsilon_r(t)$, we obtain $\varepsilon_r(t) = \frac{\varepsilon_s(t)}{\varepsilon_s(t) + (1 - \varepsilon_s(t))n^h}$. Because the protease inhibitor is given frequently, we assume both ε_s and ε_r are constants, which are used in the following simulations.

Plotting $\Phi(t)$ with typical parameter values and constant drug efficacy shows that the mutant frequency increases substantially from the pre-existing low level ($< 1\%$) to $> 5\%$ within ~ 2 days following treatment (Figure 2). This is in agreement with the results in [25]. The rapid increase in mutant frequency does not necessarily mean that the drug-resistant viral variant grows this rapidly following the telaprevir treatment. In fact, taking a close look at the eigenvalues of the system, we find $\lambda_3 < \lambda_1 < \lambda_2 < \lambda_4 < 0$ (Appendix A). This implies that both wild-type and drug resistant viruses experience a two-phase decline when we assume $T = T_0$ over a short time interval following treatment (see below for a time-varying population of target cells). Furthermore, the drug-resistant strain decreases slightly more rapidly than the wild-type strain during the first phase because $\lambda_3 < \lambda_1 < 0$ and the difference between λ_1 and λ_3 is small (Appendix A) as ε_s is close to 1, whereas it decreases slightly slower than wild-type strain in the second-phase viral decline ($\lambda_2 < \lambda_4 < 0$ and $\lambda_4 - \lambda_2$ is small). An interesting result is that the duration of the first-phase viral decline of drug-resistant virus is shorter than that of wild-type virus (Figure 2). Denoting by t_s the time at which the second-phase decline of wild-type virus begins and t_r the time at which the second-phase decline of drug-resistant virus begins, we show that $t_r < t_s$ in Appendix A. Consequently, the increase of the mutant frequency following treatment is not due to the rapid growth of drug-resistant viral variant. Rather, it is due to a longer first-phase decline of wild-type virus, unveiling the pre-existing mutant variant.

2.4 Competition between the two strains during therapy

In the above, successful suppression of the pre-existing mutant virus regardless of its drug resistance level is due to the assumption that the number of susceptible target cells remains at a constant baseline level, $T_0 = c\delta/[(1 - \mu)p_s\beta_s]$, following telaprevir treatment. If we describe the dynamics of target cells as in model (2), then drug-resistant virus is able to emerge and dominate the virus population under certain conditions.

Considering model (2), we define the reproductive ratios under treatment, $\mathcal{R}'_s = (1 - \varepsilon_s)\mathcal{R}_s$ and $\mathcal{R}'_r = (1 - \varepsilon_r)\mathcal{R}_r$. Before treatment, both strains coexist, but resistant virus remains at a very

low level ($R_s > R_r > 1$). During treatment, $\mathcal{R}'_s$ becomes less than 1 because of the efficiency of the protease inhibitor in blocking production of wild-type virus. Consequently, wild-type virus is usually successfully suppressed. If mutation only confers a low level of drug resistance (ϵ_r is large), then drug-resistant virus will also be suppressed (Figure 3, left column). However, if mutation confers high-level drug resistance (ϵ_r is small), $\mathcal{R}'_r$ may be greater than 1. Therefore, the pre-existing drug-resistant virus will outcompete wild-type and dominate the virus population under this condition (Figure 3, right column).

2.5 The effect of backward mutation

In addition to forward mutation, a backward mutation may occur that restores the original viral sequence. We compare model (1) (no backward mutation) with the following model (before treatment, i.e., $\epsilon_s = \epsilon_r = 0$) including both forward and backward mutations. Here we are assuming a single nucleotide change confers resistance, such as the G→A change that mediates the V36M mutation (i.e., the codon changes from GTG to ATG [41]), so that back mutation occurs at the same rate as forward mutation.

$$\begin{aligned}\frac{d}{dt}T(t) &= s - dT - \beta_s V_s T - \beta_r V_r T, \\ \frac{d}{dt}I_s(t) &= \beta_s V_s T - \delta I_s, \\ \frac{d}{dt}I_r(t) &= \beta_r V_r T - \delta I_r, \\ \frac{d}{dt}V_s(t) &= (1-\mu)(1-\epsilon_s)p_s I_s + \mu(1-\epsilon_r)p_r I_r - cV_s, \\ \frac{d}{dt}V_r(t) &= \mu(1-\epsilon_s)p_s I_s + (1-\mu)(1-\epsilon_r)p_r I_r - cV_r.\end{aligned}\tag{4}$$

Before drug therapy ($\epsilon_s = \epsilon_r = 0$), patients have relatively constant levels of viremia which we assume corresponds to being at the infected steady state, in which the two viral strains coexist. In Appendix B, we derive the steady states and calculate the mutant frequency

before treatment. The mutant frequency can be approximated by $\Phi = \frac{\mu}{1-r+\mu(1+r)}$, which is less than the mutant frequency in the absence of backward mutation, $\Phi = \mu/(1-r)$. However, the difference between them is minuscule since μ is small. Numerical results also suggest that including backward mutation in model (1) only has minor effects on the steady state viral loads and the pre-treatment mutant frequency (Table 1).

It is interesting to study the contribution of mutation to the dynamics of virus during therapy. Supposing that wild-type and mutant virus are both at their pretreatment baseline levels, we compare virus dynamics of the model given by Eq. (4) with the model in which both forward and backward mutations are ignored ($\mu = 0$ in (4)). Figure 4 shows the dynamics of both wild-type and drug-resistant viruses during therapy. For a mutation that confers a low level of drug resistance (for example, the mutant V36M/A confers 3.5-fold resistance [23]), inclusion of mutation has a negligible effect on the dynamics of both viral strains (Figure 4, left column). Even if mutation confers high-level resistance (for example, the A156V/T mutant confers 466-fold resistance [23]), the contribution of mutation to the level of the drug-resistant viral variant is still minor. However, in this case, wild-type virus can be maintained by backward mutation at a low level rather than being completely suppressed (Figure 4, right column). These observations are not surprising because in the presence of effective therapy targeted against wild-type virus the mutation from wild-type to drug-resistant strain makes a negligible contribution to the mutant viral load since it occurs at rate $\mu(1-\epsilon_s)$. Therefore, mutations only play a minor role in the dynamics of drug resistant virus during treatment.

Without mutation ($\mu = 0$), Eq. (4) represents a standard two-strain model in which the two strains of virus compete for the same resource (susceptible target cells). Thus, the

competitive exclusion principle applies—when the drug-resistant strain has a higher fitness under treatment ($\mathcal{R}'_r > \mathcal{R}'_s$), it outcompetes the wild-type strain.

2.6 The model with hepatocyte proliferation

Hepatocyte proliferation is important in liver regeneration [42] and can compensate for loss of hepatocytes during HCV infection. It can generate new target cells and has been included in mathematical models [43, 44]. Models with proliferation can explain complex HCV RNA profiles, such as the triphasic viral decay observed during treatment of some patients [43]. Here we incorporate proliferation of both uninfected and infected hepatocytes into model (1) and study the effects on the pretreatment mutant frequency and the evolution of drug resistance during therapy. The model with hepatocyte proliferation is

$$\begin{aligned}\frac{dT}{dt} &= s + \rho_T T \left(1 - \frac{T + I_s + I_r}{T_{max}}\right) - dT - \beta_s V_s T - \beta_r V_r T, \\ \frac{dI_s}{dt} &= \beta_s V_s T + \rho_s I_s \left(1 - \frac{T + I_s + I_r}{T_{max}}\right) - \delta I_s, \\ \frac{dI_r}{dt} &= \beta_r V_r T + \rho_r I_r \left(1 - \frac{T + I_s + I_r}{T_{max}}\right) - \delta I_r, \\ \frac{dV_s}{dt} &= (1 - \mu) p_s I_s - c V_s, \\ \frac{dV_r}{dt} &= \mu p_s I_s + p_r I_r - c V_r,\end{aligned}\tag{5}$$

where uninfected hepatocytes (i.e. target cells), hepatocytes infected with wild-type virus, and hepatocytes infected with drug resistant virus can proliferate with maximum proliferation rates ρ_T , ρ_s , and ρ_r , respectively. T_{max} is the maximum level of the total hepatocyte population. It should be noted that the value of the target cell recruitment rate, s , is different from that in model (1) because of the inclusion of proliferation in the T equation. Also, $s \ll d T_{max}$ so that in the uninfected liver $T \approx T_{max}$. Another version of this model presented elsewhere [26] includes a constant population of cells in the liver that are not susceptible to infection in the density-dependent terms of Eq. (5). This does not affect the behavior of the model but does affect parameter estimates when the model is compared with data [26].

We are interested in the pretreatment mutant frequency. In Appendix C, we show that the mutant frequency is the same as that of model (1) if $\rho_s = \rho_r$, which we expect to be the case since it is unlikely that a drug resistance mutation would affect the growth rate of an infected cell.

If we ignore mutations during treatment, then the model with hepatocyte proliferation becomes

$$\begin{aligned}\frac{dT}{dt} &= s + \rho_T T \left(1 - \frac{T + I_s + I_r}{T_{max}}\right) - dT - \beta_s V_s T - \beta_r V_r T, \\ \frac{dI_s}{dt} &= \beta_s V_s T + \rho_s I_s \left(1 - \frac{T + I_s + I_r}{T_{max}}\right) - \delta I_s, \\ \frac{dI_r}{dt} &= \beta_r V_r T + \rho_r I_r \left(1 - \frac{T + I_s + I_r}{T_{max}}\right) - \delta I_r, \\ \frac{dV_s}{dt} &= (1 - \varepsilon_s) p_s I_s - c V_s, \\ \frac{dV_r}{dt} &= (1 - \varepsilon_r) p_r I_r - c V_r.\end{aligned}\tag{6}$$

This is not a standard two-strain competition model because the two strains can coexist under certain conditions. Substituting $\tilde{V}_s = (1 - \varepsilon_s) p_s / c$ and $\tilde{V}_r = (1 - \varepsilon_r) p_r / c$ into the I_s and I_r equations, respectively, we obtain

$$\left[\frac{(1-\varepsilon_s)\beta_s p_s \bar{T}}{c} + \rho_s \left(1 - \frac{\bar{T} + \bar{I}_s + \bar{I}_r}{T_{max}}\right) - \delta \right] \bar{I}_s = 0 \quad (7)$$

and

$$\left[\frac{(1-\varepsilon_r)\beta_r p_r \bar{T}}{c} + \rho_r \left(1 - \frac{\bar{T} + \bar{I}_s + \bar{I}_r}{T_{max}}\right) - \delta \right] \bar{I}_r = 0. \quad (8)$$

If $\rho_s = \rho_r$, then it is obvious that the two strains cannot coexist because $(1 - \varepsilon_s)\beta_s p_s < (1 - \varepsilon_r)\beta_r p_r$. If $\rho_s > \rho_r$, then it is possible that the two strains coexist. In this scenario, from (7) and

(8) we have $\frac{(1-\varepsilon_s)\beta_s p_s \bar{T} - c\delta}{(1-\varepsilon_r)\beta_r p_r \bar{T} - c\delta} = \frac{\rho_s}{\rho_r}$, which yields $\bar{T} = \frac{(\rho_s - \rho_r)c\delta}{(1-\varepsilon_r)\rho_s \beta_r p_r - (1-\varepsilon_s)\rho_r \beta_s p_s}$. Combining (7) or (8) with the $\bar{T}$ equation in Eq. (6), we can obtain the steady states $\bar{s}$ and $\bar{r}$. Since their expressions are complicated, we do not present them here.

Although the two strains can coexist under certain conditions, wild-type virus is in general successfully suppressed because the inhibitor telaprevir is very effective against wild-type virus. Whether drug-resistant virus will also be suppressed depends on its reproduction capacity, proliferation potential of cells infected with resistant virus, and the drug efficacy ε_r .

3 A multi-strain model and quasispecies dynamics

3.1 Model formulation

During treatment with telaprevir, mutations mainly occur at 4 positions in the HCV NS3 protease catalytic domain, i.e., at amino acids: 36, 54, 155, and 156 [23, 25]. Here we consider the mutations occurring at these 4 positions and develop a multi-strain viral dynamic model. We assume that there is no backward mutation and that the probability of a forward mutation occurring at each amino acid is identical, denoted by μ . A schematic diagram of the mutations between these viral variants is given in Figure 5. A different multi-variant viral dynamic model was used by Adiwijaya et al. [45] to quantify the antiviral response to telaprevir and the in vivo fitness of different variants in HCV patients.

The full model with 4 positions where mutation is possible, without considering backward mutations (i.e., once a position mutates it can not mutate again), is given by

$$\begin{aligned} \frac{dT}{dt}(t) &= s - dT - \sum_j \beta_j V_j T, \\ \frac{dI_j}{dt}(t) &= \beta_j V_j T - \delta I_j, \\ \frac{dV_0}{dt}(t) &= (1-\mu)^4 p_0 I_0 - cV_0, \\ \frac{dV_i}{dt}(t) &= \mu(1-\mu)^3 p_0 I_0 + (1-\mu)^3 p_i I_i - cV_i, \quad i=1, 2, 3, 4, \\ \frac{dV_{ij}}{dt}(t) &= \mu^2(1-\mu)^2 p_0 I_0 + \mu(1-\mu)^2 (p_i I_i + p_j I_j) + (1-\mu)^2 p_{ij} I_{ij} - cV_{ij}, \quad i, j=1, 2, 3, 4 \text{ and } i < j, \\ \frac{dV_{ijk}}{dt}(t) &= \mu^3(1-\mu) p_0 I_0 + \mu^2(1-\mu) (p_i I_i + p_j I_j + p_k I_k) + \mu(1-\mu) (p_{ij} I_{ij} + p_{ik} I_{ik} + p_{jk} I_{jk}) + (1-\mu) p_{ijk} I_{ijk} - cV_{ijk}, \quad i, j, k=1, 2, 3, 4 \text{ and } i < j < k, \\ \frac{dV_{1234}}{dt}(t) &= \mu^4 p_0 I_0 + \mu^3 \sum_{i=1}^4 p_i I_i + \mu^2 \sum_{i < j} p_{ij} I_{ij} + \mu \sum_{i < j < k} p_{ijk} I_{ijk} + p_{1234} I_{1234} - cV_{1234}. \end{aligned}$$

In the first two equations, the strain index j is in the set Ω ($j \in \Omega$), where $\Omega = \{0, 1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234\}$. Strain 0 represents wild-type virus; strains 1,

2, 3, 4 represent the viral strains with mutations occurring at positions 36, 54, 155, and 156, respectively. Strains ij , $i, j = 1, 2, 3, 4$ and $i < j$, are the strains with double mutations occurring at positions i and j . Strains ijk , $i, j, k = 1, 2, 3, 4$ and $i < j < k$, and strain 1234 can be defined similarly (Figure 5). In this model, there is conservation of viruses as they mutate within the self-contained system of 16 strains.

The basic reproductive ratio for each strain is $\mathcal{R}_i = \beta_i p_i s / (dc\delta)$, $i \in \Omega$, and we define the ratio $r_i = \mathcal{R}_i / \mathcal{R}_0$, $i \in \Omega \setminus \{0\}$. r_i represents the relative fitness between mutant and wild-type virus. In the absence of selective drug pressure, r_i falls within the interval $[0, 1]$.

3.2 The frequency of the pre-existing viral variants before therapy

A tedious but straightforward calculation yields the mutant frequency of the pre-existing viral variants before treatment. The viral load of each strain is

$$\begin{aligned} V_i &= \frac{\mu}{1-\mu-r_i} V_0, \quad i=1, 2, 3, 4, \\ V_{ij} &= \frac{1}{(1-\mu)^2-r_{ij}} \left[\mu^2 V_0 + \mu(r_i V_i + r_j V_j) \right], \quad i, j=1, 2, 3, 4 \text{ and } i < j, \\ V_{ijk} &= \frac{1}{(1-\mu)^3-r_{ijk}} \left[\mu^3 V_0 + \mu^2(r_i V_i + r_j V_j + r_k V_k) + \mu(r_{ij} V_{ij} + r_{ik} V_{ik} + r_{jk} V_{jk}) \right], \quad i, j, k=1, 2, 3, 4 \text{ and } i < j < k, \\ V_{1234} &= \frac{1}{(1-\mu)^4-r_{1234}} \left[\mu^4 V_0 + \mu^3 \sum_{i=1}^4 r_i V_i + \mu^2 \sum_{\substack{i, j=1, 2, 3, 4 \\ i < j}} r_{ij} V_{ij} + \mu \sum_{\substack{i, j, k=1, 2, 3, 4 \\ i < j < k}} r_{ijk} V_{ijk} \right], \end{aligned} \quad (10)$$

where V_0 is the steady state of wild-type virus before treatment.

The above defines a recursive scheme, which allows us to orderly calculate the pretreatment steady states of the double-mutant variants (V_{ij}), 3-mutant variants (V_{ijk}), and 4-mutant variant (V_{1234}):

$$\begin{aligned} V_{ij} &= \frac{\mu^2}{(1-\mu)^2-r_{ij}} \left[1 + \frac{r_i}{1-\mu-r_i} + \frac{r_j}{1-\mu-r_j} \right] V_0, \quad i, j=1, 2, 3, 4 \text{ and } i < j, \\ V_{ijk} &= \frac{\mu^3}{(1-\mu)^3-r_{ijk}} \left[1 + \sum_{l=i, j, k} \frac{r_l}{1-\mu-r_l} + \sum_{\substack{l, m=i, j, k \\ l < m}} \frac{r_{lm}}{(1-\mu)^2-r_{lm}} \left(1 + \frac{r_l}{1-\mu-r_l} + \frac{r_m}{1-\mu-r_m} \right) \right] V_0, \quad i, j, k=1, 2, 3, 4 \text{ and } i < j < k, \\ V_{1234} &= \frac{\mu^4}{(1-\mu)^4-r_{1234}} \left[1 + \sum_{i=1}^4 \frac{r_i}{1-\mu-r_i} + \sum_{\substack{i, j=1, 2, 3, 4 \\ i < j}} \frac{r_{ij}}{(1-\mu)^2-r_{ij}} \left(1 + \frac{r_i}{1-\mu-r_i} + \frac{r_j}{1-\mu-r_j} \right) \right. \\ &\quad + \sum_{\substack{i, j, k=1, 2, 3, 4 \\ i < j < k}} \frac{r_{ijk}}{(1-\mu)^3-r_{ijk}} \left(1 + \sum_{l=i, j, k} \frac{r_l}{1-\mu-r_l} \right. \\ &\quad \left. \left. + \sum_{\substack{l, m=i, j, k \\ l < m}} \frac{r_{lm}}{(1-\mu)^2-r_{lm}} \left(1 + \frac{r_l}{1-\mu-r_l} + \frac{r_m}{1-\mu-r_m} \right) \right) \right] V_0, \end{aligned} \quad (11)$$

It follows that the pretreatment steady state level of an m -mutant viral variant is of the order of $\mu^m V_0$. The ratio of the m -mutant variant to wild-type virus, $V_{m\text{-mutant}}/V_0$, depends on the mutation rate μ , the relative fitness of the m -mutant strain and all the strains with fewer mutations. The ratio does not depend on the relative fitness of the strains with more mutations because we did not consider backward mutation in the model.

The frequency of the pre-existing viral strain i ($i \in \Omega$) before treatment is then $\Phi_i = V_i/V_{\text{total}}$

where $V_{\text{total}} = \sum_{i \in \Omega} V_i$. The frequency Φ_i depends on the mutation rate μ and the relative fitness of all mutant strains, r_i , $i \in \Omega \setminus \{0\}$.

The above formulation of the multi-strain model and the calculation of the pre-existing mutant frequency can be extended to include n possible mutations without difficulty. For example, suppose that mutations occur mainly at n positions and an m -mutant variant, $V_{m\text{-mutant}}$, has mutations occurring at the first m positions. Then the pre-existing steady state of this strain is

$$V_{m\text{-mutant}} = \frac{1}{(1-\mu)^m - r_{m\text{-mutant}}} \left[\mu^m V_0 + \mu^{m-1} \sum_{i=1}^m r_i V_i + \mu^{m-2} \sum_{\substack{i,j=1,\dots,m \\ i < j}} r_{ij} V_{ij} + \mu^{m-3} \sum_{\substack{i,j,k=1,\dots,m \\ i < j < k}} r_{ijk} V_{ijk} + \dots \right].$$

In the above expression, the steady states of the other viral variants with fewer mutations can be obtained recursively as in the scheme of equation (10). In this way, we can calculate the frequency of all pre-existing variants in a general model with n possible mutations.

The multi-strain model (Eq. 9) includes all possible viral strains that bear mutations at the four positions. However, only a few mutant strains were detected during treatment with telaprevir [23–25]. Specifically, of all the strains with two or more mutations, only strain 13 (36/155) and strain 14 (36/156) were frequently observed. Considering only these observed strains, the strain index set Ω becomes $\Theta = \{0, 1, 2, 3, 4, 13, 14\}$. Setting the appropriate r_i to 0 in Eqs. (10) and (11), the pretreatment steady states of the mutant strains are

$$\begin{aligned} V_i &= \frac{\mu}{1-\mu-r_i} V_0, \quad i=1, 2, 3, 4, \\ V_{13} &= \frac{\mu^2}{(1-\mu)^2 - r_{13}} \left(1 + \frac{r_1}{1-\mu-r_1} + \frac{r_3}{1-\mu-r_3} \right) V_0, \\ V_{14} &= \frac{\mu^2}{(1-\mu)^2 - r_{14}} \left(1 + \frac{r_1}{1-\mu-r_1} + \frac{r_4}{1-\mu-r_4} \right) V_0, \end{aligned} \quad (12)$$

where V_0 is the pretreatment steady state of the wild-type virus and $r_i = \mathcal{R}_i/\mathcal{R}_0$, $i \in \Theta \setminus \{0\}$, represents the relative fitness of strain i .

Using the in vitro estimates of the relative fitness (assuming the fitness of wild-type virus is 1) of each mutant strain in [23], we obtain the mutant frequency of the pre-existing viral variants before therapy (Table 2). For the single-mutant variant, the frequency is determined by its relative fitness—the larger the relative fitness, the higher the frequency. For the strain with two mutations, the frequency also relies on the relative fitness of those single-mutant strains that can mutate to the double-mutant strain. Although various mutant variants may exist before drug treatment [36], they only account for a very small fraction of the entire virus population. New technologies such as pyrosequencing [46] may allow one to

determine the frequency of rare mutants, but to our knowledge this has not been done for the patients in the study we have analyzed [25].

3.3 Quasispecies dynamics during therapy

As shown in previous sections, mutation during treatment has a minor effect on the dynamics of viral variants. Mutation is capable of generating low levels of viral variants that would otherwise be completely suppressed due to their fitness disadvantage, but mutation itself cannot determine which strain(s) will dominate the virus population during therapy. Here we neglect all the mutations generated during therapy. Then the multi-strain model under treatment becomes

$$\begin{aligned}\frac{d}{dt}T(t) &= s - dT - \sum_i \beta_i V_i T, \\ \frac{d}{dt}I_i(t) &= \beta_i V_i T - \delta I_i, \\ \frac{d}{dt}V_i(t) &= (1 - \varepsilon_i) p_i I_i - c V_i,\end{aligned}\tag{13}$$

where $i \in \Theta = \{0, 1, 2, 3, 4, 13, 14\}$.

The reproductive ratio for each strain under therapy is: $\mathcal{R}_i' = [(1 - \varepsilon_i) \beta_i p_i s] / (dc\delta)$, $i \in \Theta$. The above model represents a multi-strain competition system. The competitive exclusion principle also applies here. Any two viral strains cannot coexist ultimately unless they have the same reproductive ratio. Furthermore, only the viral strain with the largest reproductive ratio will persist during therapy. All the other strains with lower reproductive ratios will die out. Mathematically, if $\mathcal{R}_i' > \max_{j \in \Theta, j \neq i} (\mathcal{R}_j')$, for any $j \in \Theta$ and $j \neq i$, then the solution of the system will converge to the steady state E_i in which only strain i is present. If two or more strains have the same reproductive ratio, then they can coexist ultimately during therapy if their reproductive ratios are greater than 1 and also greater than the reproductive ratios of other strains.

4 Conclusions and discussions

Much of the recent HCV drug discovery effort has been focused on generating new therapies for HCV genotype 1 infection because of its prevalence and relatively poor response to therapy with PEG-IFN and RBV. Direct-acting antiviral agents for HCV have been suggested to be an attractive strategy whose objective is to achieve a greater response rate, with shorter treatment duration and better tolerability. However, the development of drug resistance has been a major limitation for such treatment options. The high HCV replication rate and the error-prone nature of viral RNA polymerases generate a large number of mutant viral variants, termed a quasispecies, from which variants resistant to specific drugs can be selected during treatment. Since these drug-resistant viral variants have a fitness advantage against wild-type virus in the presence of drug pressure, they are able to evolve quickly and dominate the virus population.

The HCV NS3-4A serine protease is not only involved in viral polyprotein processing but also contributes to HCV persistence by helping HCV escape the IFN antiviral response through its ability to block retinoic acid-inducible gene I and toll-like receptor-3 signaling [47, 48]. Therefore, the NS3-4A protease has become an ideal target for the development of new anti-HCV agents. Telaprevir, a new protease inhibitor, has demonstrated substantial antiviral activity in clinical studies [13, 23–25]. Administration of telaprevir even in monotherapy resulted in ~ 4-log reduction of the plasma viral load in HCV genotype 1 infected patients after 14 days [13]. However, drug-resistant viral variants were detected at high frequencies within a few days during the dosing period. The exact mechanisms

underlying emergence of viral variants such a short time after initiation of therapy is not fully characterized.

This paper studies the prevalence of the pre-existing HCV variants and the evolution of drug resistance in patients treated with a direct-acting antiviral agent such as telaprevir. We began with a simple model including two viral strains: wild-type and drug-resistant. The host cell infected with wild-type virus can produce both wild-type virus and a small fraction of drug-resistant virus due to mutations. The two strains coexist before treatment, although drug-resistant virus only accounts for a very small proportion of the virus population. The pre-existing mutant frequency, defined as the ratio of the number of the mutant virus to the total virus before treatment, is $\Phi = \mu/(1 - r)$, which is dependent only on r , the relative fitness between drug-resistant and wild-type virus, and μ , the mutation rate. Using a simplified two-strain model, we obtained an analytical solution of the mutant frequency following treatment with telaprevir. We showed that the rapid increase in the mutant frequency during therapy may not reflect the rapid replication of the pre-existing viral variants, but rather could be a consequence of the rapid and profound decline of wild-type virus, which uncovers the pre-existing mutant virus.

We studied the effects of mutation and hepatocyte proliferation on the pre-treatment mutant frequency and the evolution of drug resistance during therapy. Using the two-strain model, we showed that backward mutation has a negligible effect on the pretreatment mutant frequency. Because the protease inhibitor is highly effective against wild-type virus, both forward and backward mutations do not have a noticeable impact on the dynamics of drug-resistant virus. However, when drug-resistant virus dominates the virus population, backward mutation is able to maintain the wild-type virus at a very low level. Therefore, mutations during therapy do not contribute much to the dynamics of HCV variants. They cannot determine which strain will dominate the virus population. The dynamics of each viral strain are primarily determined by its relative fitness. Specifically, they are determined by the tradeoff between the reduced susceptibility to the protease inhibitor and resistance-associated fitness loss of the mutant virus. When hepatocyte proliferation is included in the models, the analysis becomes more complicated. In a specific case, we showed that the mutant frequency before treatment was not altered. During treatment, wild-type virus is usually suppressed by the effective agent. Whether drug-resistant virus will also be suppressed depends on the proliferation potential of cells infected with resistant virus, relative fitness of the mutant, as well as the drug efficacy. Even if the infection persists, the infected steady state can be unstable and periodic solutions may exist for an open set of parameter values [49].

We also developed a general multi-strain viral dynamic model that considers mutations among various viral strains. We derived the frequency of the pre-existing mutant variants before therapy. Without backward mutation, the frequency depends on the relative fitness of all strains that have fewer mutations. Even though including all the mutations can generate low levels of viral variants that would otherwise be significantly suppressed because of fitness disadvantage in the presence of drug pressure, the quasispecies dynamics are principally determined by the relative fitness of each strain. Mathematically, as De Leenheer and Pilyugin showed in [50], the steady state corresponding to the fittest strain in the model without mutations is globally asymptotically stable. With small perturbations due to mutations, global stability of this steady state is still preserved [50]. A few issues need also to be kept in mind when one discusses the dynamics of various viral variants during therapy. First, although only a few of all possible variants are frequently detected in clinical studies, failure to observe the other viral strains during treatment does not imply that they are not present. Indeed, even if certain viral strains are predicted to die out from the above analysis, in reality they may be present because they can be generated by mutation. However, such

strains should remain at very low levels and may exist below the detection limit of assays. Second, when we say a viral strain will die out or survive, we are referring to its steady state level (a long-term behavior). During the short dosing period of telaprevir in clinical trials, even the viral variant with the lowest fitness might be observed.

The emergence of HCV variants at high frequencies is faster than that seen in HIV. This difference can be explained by several factors related to our calculations. (1) Preexistence of drug-resistant variants. The fidelity of reverse transcriptase for HIV may be higher than that of RdRp for HCV [51, 52]. Thus, the intrinsic HCV mutation rate may be higher than that of HIV [53]. The frequency of pre-existing HCV mutants (which is proportional to μ^m for m -mutants, see calculations in Section 3.2) in the total virus population is higher than that of HIV. (2) Effectiveness of therapy. HCV protease inhibitors such as telaprevir are very effective in inhibiting the replication of wild-type virus. As we showed in Section 2.3 and Figure 2, a long and profound first-phase decline of wild-type virus uncovers the pre-existing mutant variants very fast. (3) Room for drug-resistant virus growth. Homeostatic hepatocyte proliferation can compensate for loss of liver cells during HCV infection. This generates new target cells and provides room for growth of pre-existing resistant HCV variants after therapy initiation. Some other factors such as the viral reading frames and the different turnover of the viral nucleic acid acting as a source of new viral genomes may also contribute to the rapid emergence of HCV drug resistance [54].

The prevalence of the pre-existing HCV variants has been observed in experiments. McPhee et al. [55] examined the baseline prevalence of HCV variants resistant to protease inhibitors using a highly sensitive assay (limit of detection $< 0.1\%$ of the total population). In three of eight patients, they detected the A156T variant at a frequency of 0.36%-0.75%. Cubero et al. [56] reported a similar mutant frequency (0.78%) of the A156T mutant in a chronic hepatitis C patient never treated with NS3-protease inhibitors. This frequency is higher than what we obtained in Table 2. The discrepancy can be explained either by a larger *in vivo* fitness than that in Table 2 estimated from *in vitro* experiments or by compensatory mutations (not considered in our models, see [57] that studies how compensatory mutations affect the emergence of drug resistance), which allow partial fitness recovery of the mutant variants. It has been reported that three second-site mutations, P89L, Q86R, and G162R, were able to partially reverse A156T-associated defects in polyprotein processing and/or replicon fitness without significantly reducing resistance to the protease inhibitor SCH6 [58]. In the study of Cubero et al. [56], they also detected changes at positions 89 and 86 (P89Q and Q86P) along with the A156T mutant. The presence of these mutations might compensate for the A156T-associated fitness loss and result in a higher frequency in untreated patients. The contribution of compensatory mutations to the preexistence of HCV variants and the evolution of drug resistance during treatment requires more *in vitro* and *in vivo* studies.

Our calculations may have implications for developing treatment strategies for HCV infection. From expression (11), the steady state viral level of an m -mutant strain is of the order of μ^m , which implies that a mutant variant will have a very low frequency if it carries more mutations conferring resistance to multiple drugs. In fact, in clinical trials, HCV viral variants with three or more drug-resistant mutations have seldom been identified so far. This raises the chance of success of a strategy that combines several specific HCV inhibitors targeting different steps of the HCV life cycle. The combination treatment strategy is, in theory, the same as for hepatitis B virus (HBV) [59] and HIV treatment [60]. This idea has been recently confirmed in *in vitro* studies [61–63]. When replicon cells were treated with a nucleoside HCV polymerase inhibitor in combination with either HCV-796, a non-nucleoside polymerase inhibitor, or telaprevir, the number of drug-resistant viruses was largely reduced [61], suggestive of a lack of cross resistance among the evaluated inhibitors. The data from a chimpanzee model of chronic HCV infection in which one chimp treated

with HCV protease and polymerase inhibitors was cured also support further investigation of combination therapy using only direct-acting antiviral agents [64]. Recent clinical trial results have shown that in principle patients can be cured using combinations of direct-acting antiviral agents [65, 66]. Therefore, combination of direct-acting antiviral drugs might one day be the standard therapy for HCV patients.

By calculating the fraction of all possible mutants produced per day, we estimated the number of mutations a combination of direct antivirals would need to overcome to be successful [26]. More clinical data on toxicity and drug-drug-interactions are needed in order to design combination therapies. In addition, *in vitro* data indicate that telaprevir and IFN act synergistically to inhibit HCV RNA replication and facilitate viral RNA clearance in replicon cells [67]. In clinical studies [25, 68], telaprevir was combined with PEG-IFN-alpha-2a and caused a continued antiviral response during the dosing period. More recent clinical trials [14, 15] showed that treatment with a telaprevir-based regimen significantly improved the SVR rate in patients with genotype 1 HCV. Even in patients with viral breakthrough following telaprevir alone, follow-up treatment with PEG-IFN-alpha-2a and RBV could inhibit growth of both wild-type and resistant variants [25]. These results suggest that HCV variants with reduced sensitivity to telaprevir may remain sensitive to IFN plus RBV. Based on this, both telaprevir and boceprevir were only approved by the FDA for use in combination with PEG-IFN and RBV. If we assume that IFN lowers the viral production rate by a factor $(1 - \varepsilon_{IFN})$, then the model including combination therapy of IFN and a protease inhibitor is the same as the model with the protease inhibitor monotherapy except that ε_r and ε_s are replaced with ε_{total}^s and ε_{total}^r , respectively, where $\varepsilon_{total}^s = 1 - (1 - \varepsilon_s)(1 - \varepsilon_{IFN})$ and $\varepsilon_{total}^r = 1 - (1 - \varepsilon_r)(1 - \varepsilon_{IFN})$ represent the total drug effectiveness against drug-sensitive and drug-resistant virus, respectively. Thus, the theory presented in this study can be used to also analyze combination therapy data.

Acknowledgments

Portions of this work were done under the auspices of the US Department of Energy under contract DE-AC52-06NA25396, and supported by NIH grants P30-EB011339, P20-RR018754, AI028433, OD011095, and NSF grant DMS-1122290. We also thank the reviewers for their comments that improved the manuscript.

References

1. World Health Organization. Hepatitis C. Fact sheet No 164. Revised June 2011. <http://www.who.int/mediacentre/factsheets/fs164/en/index.html>
2. Foster GR. Past, present, and future hepatitis C treatments. *Semin Liver Dis.* 2004; 24(Suppl 2):97–104. [PubMed: 15346252]
3. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, et al. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med.* 1998; 339:1485–1492. [PubMed: 9819446]
4. Davis GL, Esteban-Mur R, Rustgi V, Hoefs J, Gordon SC, et al. Interferon alfa-2b alone or in combination with ribavirin for the treatment of relapse of chronic hepatitis C. International Hepatitis Interventional Therapy Group. *N Engl J Med.* 1998; 339:1493–1499. [PubMed: 9819447]
5. Fusco DN, Chung RT. Novel therapies for hepatitis C: insights from the structure of the virus. *Annu Rev Med.* 2012; 63:373–387. [PubMed: 21942423]
6. Dore GJ, Matthews GV, Rockstroh J. Future of hepatitis C therapy: development of direct-acting antivirals. *Curr Opin HIV AIDS.* 2011; 6:508–513. [PubMed: 21897228]
7. Jazwinski AB, Muir AJ. Direct-acting antiviral medications for chronic hepatitis C virus infection. *Gastroenterol Hepatol (N Y).* 2011; 7:154–162. [PubMed: 21528041]
8. Lamarre D, Anderson PC, Bailey M, Beaulieu P, Bolger G, et al. An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature.* 2003; 426:186–189. [PubMed: 14578911]

9. Vanwolleghem T, Meuleman P, Libbrecht L, Roskams T, De Vos R, et al. Ultra-rapid cardiotoxicity of the hepatitis C virus protease inhibitor BILN 2061 in the urokinase-type plasminogen activator mouse. *Gastroenterology*. 2007; 133:1144–1155. [PubMed: 17919490]
10. Sarrazin C, Rouzier R, Wagner F, Forestier N, Larrey D, et al. SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon alpha-2b for genotype 1 nonresponders. *Gastroenterology*. 2007; 132:1270–1278. [PubMed: 17408662]
11. Lin K, Perni RB, Kwong AD, Lin C. VX-950, a novel hepatitis C virus (HCV) NS3-4A protease inhibitor, exhibits potent antiviral activities in HCV replicon cells. *Antimicrob Agents Chemother*. 2006; 50:1813–1822. [PubMed: 16641454]
12. Perni RB, Almquist SJ, Byrn RA, Chandorkar G, Chaturvedi PR, et al. Preclinical profile of VX-950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3-4A serine protease. *Antimicrob Agents Chemother*. 2006; 50:899–909. [PubMed: 16495249]
13. Reesink HW, Zeuzem S, Weegink CJ, Forestier N, van Vliet A, et al. Rapid decline of viral RNA in hepatitis C patients treated with VX-950: a phase Ib, placebo-controlled, randomized study. *Gastroenterology*. 2006; 131:997–1002. [PubMed: 17030169]
14. Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med*. 2009; 360:1839–1850. [PubMed: 19403903]
15. McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N Engl J Med*. 2009; 360:1827–1838. [PubMed: 19403902]
16. Sherman KE, Flamm SL, Afdhal NH, Nelson DR, Sulkowski MS, et al. Response-guided telaprevir combination treatment for hepatitis C virus infection. *N Engl J Med*. 2011; 365:1014–1024. [PubMed: 21916639]
17. Bacon BR, Gordon SC, Lawitz E, Marcellin P, Vierling JM, et al. Boceprevir for previously treated chronic HCV genotype 1 infection. *N Engl J Med*. 2011; 364:1207–1217. [PubMed: 21449784]
18. Poordad F, McCone JJ, Bacon BR, Bruno S, Manns MP, et al. Boceprevir for untreated chronic HCV genotype 1 infection. *N Engl J Med*. 2011; 364:1195–1206. [PubMed: 21449783]
19. Pawlotsky JM. Hepatitis C virus population dynamics during infection. *Curr Top Microbiol Immunol*. 2006; 299:261–284. [PubMed: 16568902]
20. Lin C, Gates CA, Rao BG, Brennan DL, Fulghum JR, et al. In vitro studies of cross-resistance mutations against two hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061. *J Biol Chem*. 2005; 280:36784–36791. [PubMed: 16087668]
21. Lu L, Pilot-Matias TJ, Stewart KD, Randolph JT, Pithawalla R, et al. Mutations conferring resistance to a potent hepatitis C virus serine protease inhibitor in vitro. *Antimicrob Agents Chemother*. 2004; 48:2260–2266. [PubMed: 15155230]
22. Tong X, Chase R, Skelton A, Chen T, Wright-Minogue J, et al. Identification and analysis of fitness of resistance mutations against the HCV protease inhibitor SCH 503034. *Antiviral Res*. 2006; 70:28–38. [PubMed: 16448708]
23. Sarrazin C, Kieffer TL, Bartels D, Hanzelka B, Muh U, et al. Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology*. 2007; 132:1767–1777. [PubMed: 17484874]
24. Forestier N, Reesink HW, Weegink CJ, McNair L, Kieffer TL, et al. Antiviral activity of telaprevir (VX-950) and peginterferon alfa-2a in patients with hepatitis C. *Hepatology*. 2007; 46:640–648. [PubMed: 17879366]
25. Kieffer TL, Sarrazin C, Miller JS, Welker MW, Forestier N, et al. Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis C virus replication in patients. *Hepatology*. 2007; 46:631–639. [PubMed: 17680654]
26. Rong L, Dahari H, Ribeiro RM, Perelson AS. Rapid emergence of protease inhibitor resistance in hepatitis C virus. *Sci Transl Med*. 2010; 2:30ra32.
27. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, et al. Binding of hepatitis C virus to CD81. *Science*. 1998; 282:938–941. [PubMed: 9794763]

28. Scarselli E, Ansuini H, Cerino R, Roccasecca RM, Acali S, et al. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *Embo J*. 2002; 21:5017–5025. [PubMed: 12356718]
29. Evans MJ, von Hahn T, Tscherne DM, Syder AJ, Panis M, et al. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature*. 2007; 446:801–805. [PubMed: 17325668]
30. Ploss A, Evans MJ, Gaysinskaya VA, Panis M, You H, et al. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature*. 2009; 457:882–886. [PubMed: 19182773]
31. Egger D, Wolk B, Gosert R, Bianchi L, Blum HE, et al. Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J Virol*. 2002; 76:5974–5984. [PubMed: 12021330]
32. Bartenschlager R, Frese M, Pietschmann T. Novel insights into hepatitis C virus replication and persistence. *Adv Virus Res*. 2004; 63:71–180. [PubMed: 15530561]
33. Domingo E. Biological significance of viral quasispecies. *Viral Hepatitis Rev*. 1996; 2:247–261.
34. Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science*. 1998; 282:103–107. [PubMed: 9756471]
35. Rong L, Feng Z, Perelson AS. Emergence of HIV-1 drug resistance during antiretroviral treatment. *Bull Math Biol*. 2007; 69:2027–2060. [PubMed: 17450401]
36. Bartels DJ, Zhou Y, Zhang EZ, Marcial M, Byrn RA, et al. Natural prevalence of hepatitis C virus variants with decreased sensitivity to NS3.4A protease inhibitors in treatment-naïve subjects. *J Infect Dis*. 2008; 198:800–807. [PubMed: 18637752]
37. Gillespie, J. *Population Genetics: A Concise Guide*. Johns Hopkins University Press; 1998.
38. Chu H, Herrmann E, Reesink H, Forestier N, Weegink C, et al. Pharmacokinetics of VX-950, and its effect on hepatitis C viral dynamics. *Hepatology*. 2005; 42(Suppl 1):S694.
39. Shen L, Peterson S, Sedaghat AR, McMahon MA, Callender M, et al. Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nat Med*. 2008; 14:762–766. [PubMed: 18552857]
40. Khunvichai, A.; Chu, H.; Garg, V.; McHutchison, J.; Lawitz, E., et al. Predicting HCV treatment duration with an HCV protease inhibitor co-administered with PEG-IFN/RBV by modeling both wild-type virus and low level resistant variant dynamics. *Digestive Disease Week; Washington DC*. May 19–24, 2007; 2007.
41. Zhou Y, Bartels DJ, Hanzelka BL, Muh U, Wei Y, et al. Phenotypic characterization of resistant Val36 variants of hepatitis C virus NS3-4A serine protease. *Antimicrob Agents Chemother*. 2008; 52:110–120. [PubMed: 17938182]
42. Michalopoulos GK, DeFrances MC. Liver regeneration. *Science*. 1997; 276:60–66. [PubMed: 9082986]
43. Dahari H, Ribeiro RM, Perelson AS. Triphasic decline of hepatitis C virus RNA during antiviral therapy. *Hepatology*. 2007; 46:16–21. [PubMed: 17596864]
44. Reluga TC, Dahari H, Perelson AS. Analysis of hepatitis C virus infection models with hepatocyte homeostasis. *SIAM J Appl Math*. 2009; 69:999–1023. [PubMed: 19183708]
45. Adiwijaya BS, Herrmann E, Hare B, Kieffer T, Lin C, et al. A multi-variant, viral dynamic model of genotype 1 HCV to assess the in vivo evolution of protease-inhibitor resistant variants. *PLoS Comput Biol*. 2010; 6:e1000745. [PubMed: 20419154]
46. Elahi E, Ronaghi M. Pyrosequencing: a tool for DNA sequencing analysis. *Methods Mol Biol*. 2004; 255:211–219. [PubMed: 15020827]
47. Foy E, Li K, Wang C, Sumpter JR, Ikeda M, et al. Regulation of interferon regulatory factor-3 by the hepatitis C virus serine protease. *Science*. 2003; 300:1145–1148. [PubMed: 12702807]
48. Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, et al. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature*. 2005; 437:1167–1172. [PubMed: 16177806]
49. Wang L, Li MY. Mathematical analysis of the global dynamics of a model for HIV infection of CD4+ T cells. *Math Biosci*. 2006; 200:44–57. [PubMed: 16466751]

50. De Leenheer P, Pilyugin SS. Multistrain virus dynamics with mutations: a global analysis. *Math Med Biol.* 2008; 25:285–322. [PubMed: 19015367]
51. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet.* 2008; 9:267–276. [PubMed: 18319742]
52. Powdrill MH, Tchesnokov EP, Kozak RA, Russell RS, Martin R, et al. Contribution of a mutational bias in hepatitis C virus replication to the genetic barrier in the development of drug resistance. *Proc Natl Acad Sci U S A.* 2011; 108:20509–20513. [PubMed: 22135458]
53. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. *J Virol.* 2010; 84:9733–9748. [PubMed: 20660197]
54. Soriano V, Perelson AS, Zoulim F. Why are there different dynamics in the selection of drug resistance in HIV and hepatitis B and C viruses? *J Antimicrob Chemother.* 2008; 62:1–4. [PubMed: 18441341]
55. McPhee, F.; Hernandez, D.; Zhai, G.; Friberg, J.; Yu, F., et al. Pre-existence of substitutions conferring resistance to HCV NS3 protease inhibitors. 1st International Workshop on Hepatitis C Resistance and New Compounds; Boston, MA. October 25–26, 2006;
56. Cubero M, Esteban JI, Otero T, Sauleda S, Bes M, et al. Naturally occurring NS3-protease-inhibitor resistant mutant A156T in the liver of an untreated chronic hepatitis C patient. *Virology.* 2008; 370:237–245. [PubMed: 18006035]
57. Handel A, Regoes RR, Antia R. The role of compensatory mutations in the emergence of drug resistance. *PLoS Comput Biol.* 2006; 2:e137. [PubMed: 17040124]
58. Yi M, Tong X, Skelton A, Chase R, Chen T, et al. Mutations conferring resistance to SCH6, a novel hepatitis C virus NS3/4A protease inhibitor. Reduced RNA replication fitness and partial rescue by second-site mutations. *J Biol Chem.* 2006; 281:8205–8215. [PubMed: 16352601]
59. Keeffe EB, Dieterich DT, Pawlotsky JM, Benhamou Y. Chronic hepatitis B: preventing, detecting, and managing viral resistance. *Clin Gastroenterol Hepatol.* 2008; 6:268–274. [PubMed: 18328434]
60. Yeni PG, Hammer SM, Carpenter CC, Cooper DA, Fischl MA, et al. Antiretroviral treatment for adult HIV infection in 2002: updated recommendations of the International AIDS Society-USA Panel. *JAMA.* 2002; 288:222–235. [PubMed: 12095387]
61. McCown MF, Rajyaguru S, Le Pogam S, Ali S, Jiang WR, et al. The hepatitis C virus replicon presents a higher barrier to resistance to nucleoside analogs than to nonnucleoside polymerase or protease inhibitors. *Antimicrob Agents Chemother.* 2008; 52:1604–1612. [PubMed: 18285474]
62. Wyles DL, Kaihara KA, Schooley RT. Synergy of a hepatitis C virus (HCV) NS4A antagonist in combination with HCV protease and polymerase inhibitors. *Antimicrob Agents Chemother.* 2008; 52:1862–1864. [PubMed: 18332167]
63. Wyles DL, Kaihara KA, Vaida F, Schooley RT. Synergy of small molecular inhibitors of hepatitis C virus replication directed at multiple viral targets. *J Virol.* 2007; 81:3005–3008. [PubMed: 17182685]
64. Olsen DB, Davies ME, Handt L, Koeplinger K, Zhang NR, et al. Sustained viral response in a hepatitis C virus-infected chimpanzee via a combination of direct-acting antiviral agents. *Antimicrob Agents Chemother.* 2011; 55:937–939. [PubMed: 21115793]
65. Chayama K, Takahashi S, Toyota J, Karino Y, Ikeda K, et al. Dual therapy with the nonstructural protein 5A inhibitor, daclatasvir, and the nonstructural protein 3 protease inhibitor, asunaprevir, in hepatitis C virus genotype 1b-infected null responders. *Hepatology.* 2012; 55:742–748. [PubMed: 21987462]
66. Lok AS, Gardiner DF, Lawitz E, Martorell C, Everson GT, et al. Preliminary study of two antiviral agents for hepatitis C genotype 1. *N Engl J Med.* 2012; 366:216–224. [PubMed: 22256805]
67. Lin K, Kwong AD, Lin C. Combination of a hepatitis C virus NS3-NS4A protease inhibitor and alpha interferon synergistically inhibits viral RNA replication and facilitates viral RNA clearance in replicon cells. *Antimicrob Agents Chemother.* 2004; 48:4784–4792. [PubMed: 15561857]
68. Lawitz E, Rodriguez-Torres M, Muir AJ, Kieffer TL, McNair L, et al. Antiviral effects and safety of telaprevir, peginterferon alfa-2a, and ribavirin for 28 days in hepatitis C patients. *J Hepatol.* 2008; 49:163–169. [PubMed: 18486984]

69. Powers KA, Dixit NM, Ribeiro RM, Golia P, Talal AH, et al. Modeling viral and drug kinetics: hepatitis C virus treatment with pegylated interferon alfa-2b. *Semin Liver Dis.* 2003; 23(Suppl 1): 13–18. [PubMed: 12934163]

Appendix A: Several inequalities used in the analysis of the two-strain model

1. $\lambda_3 < \lambda_1 < \lambda_2 < \lambda_4 < 0$

From $\Delta_1 = (c + \delta)^2 - 4e_s c \delta$, it is clear that $(c + \delta)^2 > \Delta_1 > (c - \delta)^2$. Thus, we have

$$\lambda_1 = -\frac{c + \delta + \sqrt{\Delta_1}}{2} < 0 \text{ and } \lambda_2 = -\frac{c + \delta - \sqrt{\Delta_1}}{2} < 0. \text{ Next we show that } \Delta_2 > \Delta_1. \text{ Calculating the}$$

difference, we obtain $\Delta_2 - \Delta_1 = 4c\delta(1 - \varepsilon_s) \left[\frac{1}{(1 - \mu)} \frac{\mathcal{R}'_r}{\mathcal{R}'_s} - 1 \right]$, where $\mathcal{R}'_r = (1 - \varepsilon_r)\mathcal{R}_r$ and $\mathcal{R}'_s = (1 - \varepsilon_s)\mathcal{R}_s$ are the reproductive ratios of the resistant and wild-type strains during therapy, respectively. Since we assume that drug-resistant virus is more fit than wild-type virus during therapy, we have that $\mathcal{R}'_r > \mathcal{R}'_s$. Thus, $\Delta_2 > \Delta_1$. Lastly, we show that $(c + \delta)^2 >$

Δ_2 . It suffices to show that $1 - \frac{(1 - \varepsilon_r)\mathcal{R}_r}{(1 - \mu)\mathcal{R}_s} > 0$, i.e., $\frac{\mathcal{R}_r}{\mathcal{R}_s} < \frac{1 - \mu}{1 - \varepsilon_r}$, which holds because wild-type virus is more fit than drug-resistant virus before treatment ($\mathcal{R}_r < \mathcal{R}_s$) and μ is very small (typically $\mu \ll \varepsilon_r$). Therefore, $(c + \delta)^2 > \Delta_2 > \Delta_1 > (c - \delta)^2$. It follows that

$$\lambda_3 = -\frac{c + \delta + \sqrt{\Delta_2}}{2} < 0 \text{ and } \lambda_4 = -\frac{c + \delta - \sqrt{\Delta_2}}{2} < 0. \text{ Furthermore, from } \sqrt{\Delta_2} > \sqrt{\Delta_1}, \text{ we have that } \lambda_3 < \lambda_1 < \lambda_2 < \lambda_4 < 0.$$

2. $C_i > 0, i = 1, 2, 3, 4$

i.

$$\text{Notice that } C_1 = -\frac{c(1 - 2\varepsilon_s) + \delta - \sqrt{\Delta_1}}{2\sqrt{\Delta_1}} V_s(0) \text{ and}$$

$\Delta_1 = (c + \delta)^2 - 4e_s c \delta$. $C_1 > 0$ is equivalent to $-c(1 - 2\varepsilon_s) - \delta + \sqrt{\Delta_1} > 0$. Thus, for $C_1 > 0$ it suffices to prove that $\sqrt{\Delta_1} > c(1 - 2\varepsilon_s) + \delta$. If the right hand side is less than 0, then the inequality automatically holds. If the right hand side is greater than 0, then we only need to show that $\Delta_1 > (c + \delta - 2c\varepsilon_s)^2$, which is equivalent to $\varepsilon_s < 1$. Hence, $\Delta_1 > (c + \delta - 2c\varepsilon_s)^2$ and $C_1 > 0$.

ii.

Because $\sqrt{\Delta_1} > c - \delta$, we have $c(1 - 2\varepsilon_s) + \delta + \sqrt{\Delta_1} > c(1 - 2\varepsilon_s) + \delta + c - \delta = 2c(1 - \varepsilon_s) > 0$.

$$\text{Thus, } C_2 = \frac{c(1 - 2\varepsilon_s) + \delta + \sqrt{\Delta_1}}{2\sqrt{\Delta_1}} V_s(0) > 0.$$

iii.

For simplicity, we introduce a new parameter θ , defined as $\theta = 1 - \frac{1 - \varepsilon_r}{1 - \mu} \frac{\mathcal{R}_r}{\mathcal{R}_s}$. Thus, $\theta < 1$. Then C_3 , C_4 , and Δ_2 can be simplified to

$$C_3 = -\frac{c(1 - 2\theta) + \delta - \sqrt{\Delta_2}}{2\sqrt{\Delta_2}} V_r(0), C_4 = \frac{c(1 - 2\theta) + \delta + \sqrt{\Delta_2}}{2\sqrt{\Delta_2}} V_r(0), \text{ and } \Delta_2 = (c + \delta)^2 - 4\theta c \delta,$$

which have similar forms to C_1 , C_2 , and Δ_1 , respectively. Following the same arguments as in (i) and (ii), we can prove that $C_3 > 0$ and $C_4 > 0$.

3. $t_r < t_s$ and t_r is an increasing function of ε_r

Since t_s is the time at which two curves $C_1 e^{\lambda_1 t}$ and $C_2 e^{\lambda_2 t}$ intersect, we obtain that

$$t_s = \frac{\ln \frac{C_1}{C_2}}{-\lambda_1 + \lambda_2} = \frac{\ln \frac{C_1}{C_2}}{\sqrt{\Delta_1}}. \text{ Similarly, we have } t_r = \frac{\ln \frac{C_3}{C_4}}{\sqrt{\Delta_2}}. \text{ Calculating the difference between } \frac{C_1}{C_2} \text{ and } \frac{C_3}{C_4}, \text{ we obtain}$$

$$\frac{C_1}{C_2} - \frac{C_3}{C_4} = \frac{-c(1-2\varepsilon_s)-\delta+\sqrt{\Delta_1}}{c(1-2\varepsilon_s)+\delta+\sqrt{\Delta_1}} - \frac{-c(1-2\theta)-\delta+\sqrt{\Delta_2}}{c(1-2\theta)+\delta+\sqrt{\Delta_2}}, \quad (14)$$

$$\text{where } \theta = 1 - \frac{1-\varepsilon_r}{1-\mu} \frac{\mathcal{R}_r}{\mathcal{R}_s}.$$

Using the common denominator to combine the two fractions in (14), we obtain the numerator

$$\left[-c(1-2\varepsilon_s)-\delta+\sqrt{\Delta_1} \right] \left[c(1-2\theta)+\delta+\sqrt{\Delta_2} \right] - \left[c(1-2\varepsilon_s)+\delta+\sqrt{\Delta_1} \right] \left[-c(1-2\theta)-\delta+\sqrt{\Delta_2} \right],$$

which can be simplified to

$$4c(\varepsilon_s \sqrt{\Delta_2} - \theta \sqrt{\Delta_1}) - 2(c+\delta)(\sqrt{\Delta_2} - \sqrt{\Delta_1}). \quad (15)$$

Because drug-resistant virus is more fit than wild-type virus during treatment, we have $(1 - \varepsilon_s) \mathcal{R}_s < (1 - \varepsilon_r) \mathcal{R}_r$. Thus, $\theta = 1 - \frac{1-\varepsilon_r}{1-\mu} \frac{\mathcal{R}_r}{\mathcal{R}_s} < \varepsilon_s$ since μ is very small. It follows from (15) that

$$4c(\varepsilon_s \sqrt{\Delta_2} - \theta \sqrt{\Delta_1}) - 2(c+\delta)(\sqrt{\Delta_2} - \sqrt{\Delta_1}) > 4c(\varepsilon_s \sqrt{\Delta_2} - \varepsilon_s \sqrt{\Delta_1}) - 2(c+\delta)(\sqrt{\Delta_2} - \sqrt{\Delta_1}) \\ = 2(\sqrt{\Delta_2} - \sqrt{\Delta_1})[2c\varepsilon_s - (c+\delta)] > 0.$$

The last inequality holds because telaprevir is very effective in blocking production of wild-type virus (ε_s is close to 1) and virus has much faster dynamics than infected hepatocytes ($c \gg \delta$). Therefore, $\frac{C_1}{C_2} > \frac{C_3}{C_4}$. Also considering that $\sqrt{\Delta_2} > \sqrt{\Delta_1}$, we have $t_s > t_r$.

Furthermore, we can prove that t_r is an increasing function with respect to ε_r , the efficacy of the protease inhibitor against the drug-resistant strain. As ε_r decreases (corresponding to a

more resistant viral strain), $\theta = 1 - \frac{1-\varepsilon_r}{1-\mu} \frac{\mathcal{R}_r}{\mathcal{R}_s}$ decreases and $\Delta_2 = (c + \delta)^2 - 4\theta c\delta$ increases.

Rearranging $\frac{C_3}{C_4}$, we have $\frac{C_3}{C_4} = \frac{-c(1-2\theta)-\delta+\sqrt{\Delta_2}}{c(1-2\theta)+\delta+\sqrt{\Delta_2}} = \frac{2\sqrt{\Delta_2}}{c(1-2\theta)+\delta+\sqrt{\Delta_2}} - 1 = \frac{2}{1+f(\theta)} - 1$, where $f(\theta) = \frac{c(1-2\theta)+\delta}{\sqrt{\Delta_2}} = \frac{c(1-2\theta)+\delta}{\sqrt{(c+\delta)^2 - 4\theta c\delta}}$. Taking derivative of $f(\theta)$ with respect to θ , we obtain

$$f'(\theta) = \frac{-2c[(c+\delta)^2 - 4\theta c\delta] + 2c\delta[c(1-2\theta) + \delta]}{[(c+\delta)^2 - 4\theta c\delta]^{\frac{3}{2}}}.$$

The numerator of the above fraction can be simplified to $2c^2(2\theta\delta - c - \delta)$, which is less than 0 because $\theta\delta < c$ and $\theta\delta < \delta$. Thus, as θ decreases, $f(\theta)$ increases. Consequently, C_3/C_4 decreases and $t_r = \ln(C_3/C_4)/\sqrt{\Delta_2}$ decreases. This shows that t_r is an increasing function of ε_r . Thus, for a mutant strain with high-level drug resistance (a small ε_r), t_r is small.

Appendix B: The mutant frequency in the model with backward mutation

There are two possible steady states of Eq. (4) before treatment: the infection-free and infected (coexistence) steady states. We are interested in the latter one. From the I_s and V_s equations, we obtain $\mu p_r \beta_r \bar{T} \bar{V}_r = [c\delta - (1-\mu)p_s \beta_s \bar{T}] \bar{V}_s$, where $\bar{T}$, $\bar{V}_s$ and $\bar{V}_r$ represent the steady states of uninfected target cells, wild-type and drug-resistant virus, respectively. Similarly, from the I_r and V_r equations, we obtain $\mu p_s \beta_s \bar{T} \bar{V}_s = [c\delta - (1-\mu)p_r \beta_r \bar{T}] \bar{V}_r$. Thus, the two strains coexist only when $(1-\mu)p_s \beta_s \bar{T} < c\delta$ and $(1-\mu)p_r \beta_r \bar{T} < c\delta$. From the above two equations, we obtain an equation that the steady state of uninfected hepatocytes, $\bar{T}$, must satisfy $(1-2\mu)p_s \beta_s p_r \beta_r \bar{T}^2 - (1-\mu)c\delta(p_s \beta_s + p_r \beta_r) \bar{T} + (c\delta)^2 = 0$, which has two solutions:

$$\bar{T}_{1,2} = \frac{(1-\mu)(p_s \beta_s + p_r \beta_r) \pm \sqrt{[(1-\mu)(p_s \beta_s + p_r \beta_r)]^2 - 4(1-2\mu)p_s \beta_s p_r \beta_r}}{2(1-2\mu)p_s \beta_s p_r \beta_r} c\delta. \quad (16)$$

Ignoring μ , we have two approximate solutions: $\bar{T}_1 \approx \frac{c\delta}{p_r \beta_r}$ (choosing “+” in (16)) and $\bar{T}_2 \approx \frac{c\delta}{p_s \beta_s}$ (choosing “-” in (16)). Because of the conditions for the existence of the coexistence steady state and the assumption that $\beta_r < \beta_s$ and $p_r < p_s$, only $\bar{T}_2$ is feasible.

Thus, the mutant frequency before treatment is $\Phi = \frac{\bar{V}_r}{\bar{V}_s + \bar{V}_r} = \frac{1}{1 + \frac{c\delta - (1-\mu)p_r \beta_r \bar{T}_2}{\mu p_s \beta_s \bar{T}_2}}$, where

$$\bar{T}_2 = \frac{(1-\mu)(p_s \beta_s + p_r \beta_r) - \sqrt{[(1-\mu)(p_s \beta_s + p_r \beta_r)]^2 - 4(1-2\mu)p_s \beta_s p_r \beta_r}}{2(1-2\mu)p_s \beta_s p_r \beta_r} c\delta.$$

Using $\bar{T}_2$, Φ can be further simplified to

$$\Phi = \frac{\mu}{\frac{(1-\mu)(1+r) + \sqrt{[(1-\mu)(1+r)]^2 - 4(1-2\mu)r}}{2} - r + \mu(1+r)}, \quad (17)$$

where $r = \beta_r / \beta_s$. It is clear that Φ depends only on μ and r .

It follows from (17) that F can be approximated by

$$\Phi = \frac{\mu}{1-r+\mu(1+r)}, \quad (18)$$

which is less than $\frac{\mu}{1-r}$, the mutant frequency in the model without considering backward mutation. In fact, it can be proved rigorously that $\Phi_w < \Phi_{wo}$, where Φ_w represents the

mutant frequency with backward mutation (defined in (17)) and $\Phi_{wo} = \frac{\mu}{1-r}$ is the mutant frequency without backward mutation. For the proof, it suffices to show that

$$\frac{(1-\mu)(1+r) + \sqrt{[(1-\mu)(1+r)]^2 - 4(1-2\mu)r}}{2} + \mu(1+r) > 1.$$
 This inequality is equivalent to $r < 1$, which holds because resistant virus is less fit than wild-type virus in the absence of treatment ($\mathcal{R}_r < \mathcal{R}_s$). Therefore, $\Phi_w < \Phi_{wo}$. However, from the approximation of Φ_w (Eq. (18)), we observe that the difference between Φ_w and Φ_{wo} is miniscule. This shows that backward mutation only plays a minor role in the pre-treatment mutant frequency.

Appendix C: The pretreatment mutant frequency in the model with hepatocyte proliferation

From the V_s and V_r equations of model (5), we have $\bar{V}_s = \frac{(1-\mu)p_s\bar{I}_s}{c}$ and $\bar{V}_r = \frac{\mu p_s\bar{I}_s + p_r\bar{I}_r}{c}$.

Substituting into the I_s and I_r equations, we obtain $\frac{(1-\mu)\beta_s p_s \bar{T}}{c} + \rho_s(1 - \frac{\bar{T} + \bar{I}_s + \bar{I}_r}{T_{max}}) = \delta$ and

$\frac{\beta_r(\mu p_s\bar{I}_s + p_r\bar{I}_r)\bar{T}}{c\bar{I}_r} + \rho_r(1 - \frac{\bar{T} + \bar{I}_s + \bar{I}_r}{T_{max}}) = \delta$. If $\rho_s = \rho_r$, from the above two equations we have

$$\frac{(1-\mu)\beta_s p_s \bar{T}}{c} = \frac{\beta_r(\mu p_s\bar{I}_s + p_r\bar{I}_r)\bar{T}}{c\bar{I}_r},$$
 which yields $\bar{I}_r = \frac{\mu\beta_r p_s \bar{I}_s}{(1-\mu)\beta_s p_s - \beta_r p_r}$. Substituting into

$\bar{V}_r = \frac{\mu p_s\bar{I}_s + p_r\bar{I}_r}{c}$, we have $\bar{V}_r = \frac{\mu p_s\bar{I}_s}{c}(1 + \frac{r}{1-\mu-r})$, where r denotes the ratio $\mathcal{R}_r / \mathcal{R}_s$.

Considering $\bar{V}_s = (1-\mu)p_s\bar{I}_s/c$, we obtain the mutant frequency $\Phi = \frac{\bar{V}_r}{\bar{V}_r + \bar{V}_s} = \frac{\mu}{1-r}$, which is the same as the mutant frequency in the model without hepatocyte proliferation. It should be noted that although the mutant frequency is the same as that in the model without hepatocyte proliferation, the steady states of wild-type and resistant virus are not necessarily the same as the previous ones. They depend on ρ_T , ρ_s , ρ_r , and other parameters.

Figure 1. HCV life cycle

(1) Following viral binding, receptor-mediated endocytosis and membrane fusion, nucleocapsid enters into the cytoplasm of the host cell; (2) Uncoating of nucleocapsid exposes a positive-strand RNA genome; (3) Internal ribosomal entry site (IRES)-mediated translation of the viral genome generates a large polyprotein, which is then proteolytically cleaved by enzymes such as the NS3-4A serine protease to produce 10 viral proteins; (4) Viral polymerase, a product of cleavage, participates in the synthesis of both positive- and negative-strand RNA genomes; (5) Packaging and assembly of progeny virions; (6) Vesicle fusion at the plasma membrane and viral release.

Figure 2. Mutant frequency and virus dynamics after treatment

The target cell level is assumed to be the pre-treatment steady state T_0 . Thick solid line is the viral load and dashed line is the mutant frequency after therapy. t_s is the time at which the second-phase decline of wild-type virus begins, and t_r is the time at which the second-phase decline of drug resistant virus begins. Model parameters are: $c = 6.2 \text{ day}^{-1}$ [34], $\delta = 0.14 \text{ day}^{-1}$ [34], $\mu = 10^{-4}$ per copied nucleotide [33], $e_s = 0.9997$ [38], the Hill coefficient is $h = 2$ [69]. We assumed the mutant, for example T54A, confers 12-fold resistance and $\mathcal{R}_r / \mathcal{R}_s = 0.81$ [23]. We obtained the eigenvalues $\lambda_1 = -6.2$, $\lambda_2 = -0.14$, $\lambda_3 = -6.2048$, $\lambda_4 = -0.1352$, and $t_s = 1.33$ day, $t_r = 0.55$ day.

Figure 3. Competition between wild-type (upper panels) and drug-resistant virus (middle panels) during therapy (Eq. (2))

The total viral levels are plotted in the lower panels. Left column: assuming the mutant, for example V36A/M, confers 3.5-fold resistance and $\mathcal{R}_r/\mathcal{R}_s=0.98$ [23]. Both wild-type and drug-resistant virus are suppressed. Right column: assuming the mutant, for example A156V/T, confers 466-fold resistance and $\mathcal{R}_r/\mathcal{R}_s=0.45$ [23]. Wild-type virus is suppressed, whereas drug-resistant virus arises and dominates the virus population, which results in a viral rebound in the total viral level. The values of parameters used are [26, 34]: $s = 7.5 \times 10^5 \text{ cells mL}^{-1} \text{ day}^{-1}$, $d = 0.01 \text{ day}^{-1}$, $\beta_s = \beta_r = 10^{-7} \text{ mL day}^{-1} \text{ virions}^{-1}$, $\mu = 10^{-4} \text{ per copied nucleotide}$, $c = 6.2 \text{ day}^{-1}$, $\delta = 0.14 \text{ day}^{-1}$, $p_s = 10 \text{ virions cell}^{-1} \text{ day}^{-1}$, $e_s = 0.9997$, and the Hill coefficient is $h = 2$.

Figure 4. Contribution of mutation to the dynamics of wild-type and drug-resistant virus during treatment

Assuming there is no mutation during treatment (thick dashed line) and there exist both forward and backward mutations during therapy (thin solid line). Left column: assuming the mutant confers 3.5-fold resistance and $\mathcal{R}_r/\mathcal{R}_s=0.98$ [23]. The solid and the dashed lines almost overlap, which suggests that mutation has a negligible effect on the evolution of both strains when the mutation confers a low level of drug resistance. Right column: assuming the mutant confers 466-fold resistance and $\mathcal{R}_r/\mathcal{R}_s=0.45$ [23]. Mutation still does not contribute much to the evolution of drug-resistant virus, which emerges and dominates the virus population. However, wild-type virus is maintained at a low level by backward mutation rather than being completely suppressed. The values of parameters used are [26, 34]: $s = 7.5 \times 10^5 \text{ cells mL}^{-1} \text{ day}^{-1}$, $d = 0.01 \text{ day}^{-1}$, $\beta_s = \beta_r = 10^{-7} \text{ mL day}^{-1} \text{ virions}^{-1}$, $\mu = 10^{-4} \text{ per copied nucleotide}$, $c = 6.2 \text{ day}^{-1}$, $\delta = 0.14 \text{ day}^{-1}$, $p_s = 10 \text{ virions cell}^{-1} \text{ day}^{-1}$, $\epsilon_s = 0.9997$, and $h = 2$.

Figure 5. Mutation diagram

From the topmost (layer 0) to the bottommost (layer 4) are the wild-type, single-mutant, double-mutant, 3-mutant, and 4-mutant strains. Backward mutation is not considered. Mutations between strains that lead to transitions beyond one layer are not shown but are considered in the multi-strain model. For example, the mutation from the wild-type (layer 0) to strain 12 (layer 2) or strain 123 (layer 3) is not plotted in the diagram.

Table 1Effects of backward mutation on the steady states and the mutant frequency before treatment^{*}

Model	Steady state of the wild-type virus V_s	Steady state of the resistant virus V_r	Mutant frequency $\Phi = V_r / (V_s + V_r)$
Without backward mutation (Eq. (1))	4.9386×10^6 IU/mL	1.2350×10^3 IU/mL	2.5000×10^{-4}
With backward mutation (Eq. (4))	4.9386×10^6 IU/mL	1.2348×10^3 IU/mL	2.4996×10^{-4}

^{*} Model parameters used to obtain the steady states of V_s and V_r are [26, 34]: $d = 0.01 \text{ day}^{-1}$, $p_s = 10 \text{ virions cell}^{-1} \text{ day}^{-1}$, $\mu = 10^{-4}$, $c = 6.2 \text{ day}^{-1}$, $\delta = 0.24 \text{ day}^{-1}$. Assuming $T(0)$ is about 1.5×10^6 cells/mL and $V_s(0)$ is about 5×10^6 IU/mL at the baseline before treatment, we have $\beta_s = 10^{-7} \text{ mL day}^{-1} \text{ virions}^{-1}$ and $s = 7.5 \times 10^5 \text{ cells mL}^{-1} \text{ day}^{-1}$. For simplicity, we assume that wild-type and resistant viruses differ only in their replication capacities. We choose $\beta_r = \beta_s = 10^{-7} \text{ mL day}^{-1} \text{ virions}^{-1}$ and $p_r = 6 \text{ virions cell}^{-1} \text{ day}^{-1}$ (supposing that a single-mutant variant, for example R155K/T, confers ~10-fold resistance and has a relative fitness of ~0.6 compared with wild-type virus [23]).

Table 2Estimated mutant frequency of the pre-existing viral variants before therapy^{*}

Mutant viral variants	Relative fitness ^{**}	Pretreatment frequency
V36A/M	0.98	5.00×10^{-3}
T54A	0.81	5.23×10^{-4}
R155K/T	0.62	2.62×10^{-4}
A156V/T	0.45	1.81×10^{-4}
36/155	0.82	2.87×10^{-6}
36/156	0.67	1.54×10^{-6}

^{*} A typical HCV infected patient has a viral load $> 10^6$ per ml or $> 10^{10}$ virions in the total extracellular body water. Thus, the total number of each pre-existing viral variant is sufficiently large ($> 10^4$ virions) that prediction via a deterministic model is justified.

^{**} Relative fitness is from the in vitro estimates in [23].