LA-UR- 08-4751

Approved for public release; distribution is unlimited.

Title: The Prevalence of the Pre-Existing Hepatitis C Viral Variants and the Evolution of Drug Resistance in Patients Treated with the NS3-4A Serine Protease Inhibitor Telaprevir.

Author(s): L. Rong, Z # 216571, T-10/T-Division
R.M. Ribeiro, Z # 171295, T-10/T-Division
A. Perelson, Z # 082011, T-10/T-Division

Intended for: Journal- PLoS Computational Biology

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

The Prevalence of the Pre-Existing Hepatitis C Viral Variants and the Evolution of Drug Resistance in Patients Treated with the NS3-4A Serine Protease Inhibitor Telaprevir

Libin Rong
Theoretical Biology and Biophysics
Los Alamos National Laboratory, Los Alamos, NM 87545

Ruy M. Ribeiro
Theoretical Biology and Biophysics
Los Alamos National Laboratory, Los Alamos. NM 87545

Alan S. Perelson
Theoretical Biology and Biophysics
Los Alamos National Laboratory, Los Alamos, NM 87545

Running title: HCV Drug Resistance to Telaprevir

Corresponding author:
Dr. Alan Perelson
Mailing Address: MS K710, Los Alamos National Laboratory, Los Alamos, NM 87545.
Phone: 505-667-6829; Fax: 505-665-3493; Email: asp@lanl.gov

Author contributions:

LR and ASP conceived and designed the experiments. LR performed the experiments. LR analyzed the data. LR contributed reagents/materials/analysis tools. RMR and ASP advised on model structure and analysis. All authors wrote the paper.

Word count:

Abstract: 292; Author summary: 185; Text: 7484

Abstract

Telaprevir (VX-950), a novel hepatitis C virus (HCV) NS3-4A serine protease inhibitor, has demonstrated substantial antiviral activity in patients infected with HCV genotype 1. Some patients experience viral breakthrough, which has been shown to be associated with emergence of telaprevir-resistant HCV variants during treatment. The exact mechanisms underlying the rapid selection of drug resistant viral variants during dosing are not fully understood. In this paper, we develop a two-strain model to study the pre-treatment prevalence of the mutant virus and derive an analytical solution of the mutant frequency after administration of the protease inhibitor. Our analysis suggests that the rapid increase of the mutant frequency during therapy is not due to mutant growth but rather due to the rapid and profound loss of wild-type virus, which uncovers the pre-existing mutant variants. We examine the effects of backward mutation and hepatocyte proliferation on the pre-existence of the mutant virus and the competition between wild-type and drug resistant virus during therapy. We then extend the simple model to a general model with multiple viral strains. Mutations during therapy do not play a significant role in 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. Hepatocyte 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. Our study provides a theoretical framework for exploring the prevalence of pre-existing mutant variants and the evolution of drug resistance during treatment with other protease inhibitors or HCV polymerase inhibitors.

Author Summary

Chronic infection with hepatitis C virus (HCV) still remains an important health-care problem worldwide despite significant progress in the development of HCV treatment since the discovery of the virus in 1989. The current standard therapy is effective only in a fraction of treated patients. Telaprevir, a new HCV protease inhibitor, has demonstrated a promising result in clinical studies. However, drug resistant HCV variants were detected in the population of virus a few days after drug administration. We have developed a mathematical model that can explain the rapid selection of drug resistance in HCV patients treated with telaprevir. We explored the potential influences of backward mutation and target cell proliferation on the dynamics of both drug sensitive and resistant viruses. By developing a multi-strain viral dynamic model, we further studied the pretreatment frequency of viral variants and HCV quasispecies dynamics during therapy. Our work provides a mathematical framework that can be employed to study the preexistence and the evolution of drug resistance in HCV patients treated with other protease inhibitors and HCV polymerase inhibitors, and may have significant implications for the treatment of HCV infection.

Introduction

Chronic viral infection with hepatitis C virus (HCV) has caused an epidemic with approximately 170 million people infected worldwide and 3 to 4 million individuals newly infected each year [1]. About 80% of newly infected patients progress to develop chronic infection [2]. Of those chronically infected, a proportion of patients develop serious liver diseases such as cirrhosis and hepatocellular carcinoma [3]. The current standard therapy for HCV infection consists of pegylated interferon (PEG-IFN) administered once weekly, along with daily oral ribavirin (RBV) for 24 or 48 weeks [4–6]. Although the combination exerts synergistic antiviral effects [7], it leads to sustained elimination of the virus in only some treated patients. The HCV genotype appears to be the most important factor in predicting response. In patients infected with genotypes 2 and 3, about 90% of patients achieve sustained viral response (SVR), defined as the absence of detectable serum HCV RNA 24 weeks after completion of treatment; whereas in patients infected with genotype 1, the major genotype affecting North America, Europe and Japan, only about 40% of treated individuals show SVR [5]. 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 inhibitors that target different steps of the HCV life cycle. Such antiviral agents represent the concept of specifically targeted antiviral therapy for HCV (STAT-C) (see reviews in [8]). An important target is the HCV-encoded NS3-4A serine protease [9]. In clinical trials HCV protease inhibitors have been tested 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 [10], but clinical development was halted due to drug-induced cardiotoxicity [11, 12]. SCH 503034 (boceprevir; Schering-Plough Pharmaceuticals), another oral HCV protease inhibitor, also demonstrated substantial antiviral effects when used in combination with PEG-IFN-alpha-2b in HCV genotype 1 infected patients who were previously nonresponders to PEG-IFN-alpha-2b +/− RBV therapy [13]. 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 [14]. It had a favorable pharmacokinetic profile with high exposure in the liver in several animal models [15], 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 [16].

Emergence of drug resistant mutations is a problem challenging the development of specifically targeted antiviral drugs (see reviews in [8, 17]). 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 [18–21], whose composition is subject to continuous change due to the competition between newly generated mutants and existing variants with different phenotypes and fitness [22–24]. 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 rapidly become the dominant genotype [25, 26]. The amino acid substitutions that were reported to be selected by protease inhibitors and to confer drug resistance have been characterized *in vitro* in the HCV replicon system [27–31].

Recently, the initial selection and kinetics of telaprevir-resistant HCV variants have been described in patients given the protease inhibitor alone [16, 32] or in combination with PEG-IFN-alpha-2a [33, 34]. 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 [16]. Using a highly sensitive sequencing assay, Sarrazin et al. [32] identified mutations that confer resistance to telaprevir in the NS3 protease catalytic domain and correlated them with virologic response. These mutations were further confirmed in a subsequent study [34] that provides a detailed kinetic analysis of HCV variants in patients treated with telaprevir alone or in combination with PEG-IFN-alpha-2a for 14 days. Four of the 8 patients in the telaprevir group exhibited viral load rebound during the dosing period. Virus isolated from these patients at day 3 contained low levels (5%–20%) of single-mutant resistant variants, which increased in the population of virus isolated at days 7 and 11, and were replaced by more resistant double-mutant variants by day 14 (end of dosing) and during the first follow-up week with standard therapy [34]. Why drug resistant viral variants were selected this rapidly following treatment with telaprevir remains unclear. The potential mechanisms underlying the rapid emergence of mutant variants and the evolution of drug resistance during therapy are the subject of this study.

For human immunodeficiency virus (HIV) antiretroviral therapy, mathematical models have been developed to investigate the frequency of pre-existing mutant variants [35], emergence of

drug resistance following treatment [26, 36–38], evolution of drug resistant virus [39–41] and HIV quasispecies dynamics [42, 43] during therapy. In this study, we address these issues in HCV patients during treatment 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. Steady state analysis suggests that both strains coexist before treatment and the pre-existing mutant frequency depends on only the mutation rate and the relative viral fitness between resistant and wild-type virus. With reasonable simplifications, we develop 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 following treatment. We describe the competition between wild-type and drug resistant virus during treatment. We also study 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 in more than one mutation 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 are also provided. Our work offers a mathematical framework that can be used to study the prevalence of pre-existing mutant variants and the evolution of drug resistance during treatment with other protease inhibitors or HCV polymerase inhibitors.

Results

A Two-Strain Model

Before describing the model, we use a hypothetical HCV life cycle (Figure 1) as a framework for discussing our current knowledge of virus replication (see reviews of the HCV life cycle in [44, 45]). The exact mechanism by which HCV enters hepatocytes, the primary targets of infection, is still largely unknown. It is presumably receptor-mediated and possibly involves CD81 [46] and the human scavenger receptor class B type 1 (SR-B1) [47]. 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 non-structural (NS) proteins NS2, NS3, NS4A, NS4B, NS5A and NS5B. This is followed by HCV RNA replication that occurs in a specific cytoplasmic membrane alteration, termed the 'membranous web', whose formation is induced by the integral membrane protein NS4B [48]. The process of RNA synthesis is not fully characterized, but is likely to be semi-conservative and asymmetric [49]: 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 HCV 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 [50]. Furthermore, since the RdRp is devoid of proofreading capacity and other postreplicative repair mechanisms, it cannot correct misincorporations that occur randomly during replication [20]. The high mutation rate, together with rapid HCV replication [51] and the large viral population size, results in progressive diversification of viral genotypes and subtypes in geographically or epidemiologically populations, and in the quasispecies distribution of the virus population in a given infected individual [20].

We adapt a mathematical model, which was developed to study HIV-1 infection [52] and drug resistance [26], 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 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 has a probability μ to produce 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 [26] 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 comparisons later.

The Frequency of the Mutant Virus before Treatment

With $\mu > 0$, there are three possible steady states of model (1): the infection-free steady state (E_0), the steady state in which only drug resistant virus is present (E_r), and the steady state in which the two strains coexist (E_c) (Supporting Information 1). Defining the basic reproductive ratios [53], $\mathcal{R}_s = s\beta_s p_s/(dc\delta)$ and $\mathcal{R}_r = s\beta_r p_r/(dc\delta)$, of the wild-type and the drug resistant strains, respectively, we obtain the conditions for the existence of these steady states: E_r is feasible if and only if $\mathcal{R}_r > 1$; E_c is feasible if and only if $\mathcal{R}_s > \max(1/(1-\mu), \mathcal{R}_r/(1-\mu))$. As the basic reproductive ratio measures the number of progeny virions of the first generation produced by a single virus in a healthy host, the ratio of $\mathcal{R}_r$ to $\mathcal{R}_s$, i.e., $r = \mathcal{R}_r/\mathcal{R}_s = \beta_r p_r/(\beta_s p_s)$, represents the relative fitness of drug resistant to wild-type in the absence of drug pressure.

The above existence conditions also provide threshold conditions for the stability of the steady states. Indeed, in Supporting Information 1, we 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. Considering the predominance of wild-type virus before treatment ($\mathcal{R}_s > 1$) and resistance-associated loss of fitness ($\mathcal{R}_r < \mathcal{R}_s$, i.e., $r < 1$) [32] as in HIV-1 [54], the conditions in (iii) are typically satisfied because μ is very small. As a consequence, the solutions of model (1) converge to the steady state E_c, i.e., both wild-type and drug resistant viral strains coexist in infected individuals before therapy.

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. Using the results in Supporting Information 1, 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. This is consistent with the result of the frequency of resistant mutant HIV-1 before antiretroviral treatment [35]. 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 as above is equivalent to the result in population genetics where the mutant frequency is derived by the mutation-selection balance [55], 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).

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 [15] and has been shown to profoundly reduce the plasma viral load in infected individuals [16, 32]. 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 $0 \leq \epsilon_s,\ \epsilon_r \leq 1$ with $\epsilon = 1$ being a 100% effective drug, the model under treatment with the protease inhibitor reads

$$\begin{aligned}
\frac{d}{dt}T(t) &= s - dT - \beta_s V_s T - \beta_r V_r T,\\
\frac{d}{dt}I_s(t) &= \beta_s V_s T - \delta I_s,\\
\frac{d}{dt}I_r(t) &= \beta_r V_r T - \delta I_r,\\
\frac{d}{dt}V_s(t) &= (1-\mu)(1-\epsilon_s)p_s I_s - cV_s,\\
\frac{d}{dt}V_r(t) &= \mu(1-\epsilon_s)p_s I_s + (1-\epsilon_r)p_r I_r - cV_r.
\end{aligned} \tag{2}$$

Assuming T remains at a constant level, T_0, over a period of several days following treatment, and ignoring the term $\mu(1-\epsilon_s)p_s I_s$ in the V_r equation (because μ is $\sim 10^{-4}$–10^{-5} [50] and ϵ_s is close to 1 [56]), we can reduce (2) to a solvable system and develop an analytical solution of the mutant frequency, $\Phi(t) = V_r(t)/(V_s(t)+V_r(t))$, after drug administration (see Materials and Methods). $\Phi(t)$ depends on c, δ, μ, ϵ_s, ϵ_r, and the relative fitness r.

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 [57, 58]: $\epsilon_s(t) = C(t)^h/(\mathrm{IC}_{50}{}^h + C(t)^h)$, 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) [59], which is consistent with the 3-4 log first-phase drop of plasma HCV RNA levels when telaprevir was administered in monotherapy [16]. Similarly, for the mutant virus with n-fold resistance, i.e., an n-fold increase of IC_{50}, we have $\epsilon_r(t) = C(t)^h/[(n \cdot \mathrm{IC}_{50})^h + C(t)^h]$. From the above two equations, we obtain the drug efficacy for an n-fold resistant mutant strain based on the drug efficacy for the wild-type strain: $\epsilon_r(t) = \epsilon_s(t)/[\epsilon_s(t) + (1 - \epsilon_s(t))n^h]$.

Plotting $\Phi(t)$ with typical parameter values and constant drug efficacy (median) shows that the mutant frequency increases substantially from the pre-existing low level ($< 1\%$) to $> 5\%$ within ~ 2 days following treatment (Figure 2). However, the rapid increase of the 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 $\lambda_3 < \lambda_1 < \lambda_2 < \lambda_4 < 0$ (Supporting Information 2), which implies that both wild-type and drug resistant viruses experience a two-phase decline when we assume $T = T_0$ over a period of several days following treatment (see below). Furthermore, the drug resistant strain decreases slightly more rapidly than wild-type strain during the first phase because $\lambda_3 < \lambda_1 < 0$ and the difference between λ_1 and λ_3 is small (Supporting Information 2) 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 Supporting Information 2. 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.

Competition between the Two Strains during Therapy

Suppression of the pre-existing mutant virus regardless of its drug resistance level, as shown in Figure 2, 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 the telaprevir treatment. If we describe the dynamics

of target cells as in the model given by Eq. (2), then drug resistant virus is able to emerge and ultimately dominate the virus population under certain conditions.

Considering the model given by Eq. (2), we define the reproductive ratios under treatment, $\mathcal{R}'_s = (1-\epsilon_s)\mathcal{R}_s$ and $\mathcal{R}'_r = (1-\epsilon_r)\mathcal{R}_r$. Before treatment, both strains coexist, but resistant virus remains at a very low level ($\mathcal{R}_s > \mathcal{R}_r > 1$). 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, following the steady state analysis in Supporting Information 1, the pre-existing drug resistant virus will outcompete wild-type and dominate the virus population under this condition (Figure 3, right column). In this case, however, the increase of drug resistant virus to a high level takes more time than the increase of the mutant frequency derived in the last section (compare Figures 2c and 3e).

The Effect of Backward Mutation

We compare model (1) (no backward mutation) with the following model (before treatment, i.e., $\epsilon_s = \epsilon_r = 0$) including backward mutation from resistant to drug sensitive virus. 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), so that back mutation occurs at the same rate as forward mutation.

$$\begin{aligned}
\frac{d}{dt}T(t) &= s - dT - \beta_s V_s T - \beta_r V_r T,\\
\frac{d}{dt}I_s(t) &= \beta_s V_s T - \delta I_s,\\
\frac{d}{dt}I_r(t) &= \beta_r V_r T - \delta I_r,\\
\frac{d}{dt}V_s(t) &= (1-\mu)(1-\epsilon_s)p_s I_s + \mu(1-\epsilon_r)p_r I_r - cV_s,\\
\frac{d}{dt}V_r(t) &= \mu(1-\epsilon_s)p_s I_s + (1-\mu)(1-\epsilon_r)p_r I_r - cV_r.
\end{aligned} \tag{3}$$

Before drug therapy ($\epsilon_s = \epsilon_r = 0$), the above model remains at the infected steady state, in which the two viral strains coexist. In Supporting Information 3, we derive the steady states and calculate the mutant frequency before treatment. The mutant frequency can be approximated by $\Phi = \mu/[1 - r + \mu(1+r)]$, which is less than the mutant frequency in the absence of backward

mutation, $\Phi = \mu/(1-r)$. However, the difference between them is miniscule. Numerical results also suggest that including backward mutation in model (1) only has minor effects on the steady state viral load and the pre-treatment mutant frequency (see Table 1).

It is interesting to study the contribution of mutation to the dynamics of the pre-existing drug resistant 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. (3) with the model in which both forward and backward mutations are ignored ($\mu = 0$ in (3)).

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 [32]), 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 [32]), 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. (3) 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. On the contrary, if $\mathcal{R}_s' > \mathcal{R}_r'$, then wild-type virus dominates the virus population (see Supporting Information 1, $\mu = 0$).

The Model with Hepatocyte Proliferation

Hepatocyte proliferation, which is important in liver regeneration [60], can also compensate for loss of hepatocytes during HCV infection, and thus has been included in mathematical models [61]. Models with proliferation can explain complex HCV RNA profiles, such as the triphasic viral decay observed during treatment of some patients [62]. 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

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

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 the previous one in model (1) because of the inclusion of proliferation in the T equation. Also, $s \leq dT_{max}$ so that in the uninfected liver $T \leq T_{max}$.

We are interested in the pretreatment mutant frequency. In Supporting Information 4, 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.

Even when we ignore mutations during treatment, the model with hepatocyte proliferation is not a standard two-strain competition model because of the individual proliferation of infected hepatocytes, I_s and I_r. In fact, the two strains can coexist under certain conditions during treatment (see Materials and Methods). However, 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.

A Multi-Strain Model

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 [32, 34]. Here we consider the mutations occurring at these 4 positions and develop a multi-strain viral dynamic model. We ignore backward mutation and assume that the probability of a 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. The multi-strain model and the pretreatment frequency of drug resistant viral variants are given in Materials and Methods.

The multi-strain model includes all possible viral strains that bear mutations at the four positions. However, only a few mutant strains were detected during treatment with telaprevir [32–34]. 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 general multi-strain model is reduced to

$$\begin{aligned}
&\frac{d}{dt}T(t) = s - dT - \sum_j \beta_j V_j T,\\
&\frac{d}{dt}I_j(t) = \beta_j V_j T - \delta I_j,\\
&\frac{d}{dt}V_0(t) = (1-\mu)^4 p_0 I_0 - cV_0,\\
&\frac{d}{dt}V_i(t) = \mu(1-\mu)^3 p_0 I_0 + (1-\mu)^4 p_i I_i - cV_i, \quad i = 1,2,3,4,\\
&\frac{d}{dt}V_{13}(t) = \mu^2(1-\mu)^2 p_0 I_0 + \mu(1-\mu)^3(p_1 I_1 + p_3 I_3) + (1-\mu)^4 p_{13} I_{13} - cV_{13},\\
&\frac{d}{dt}V_{14}(t) = \mu^2(1-\mu)^2 p_0 I_0 + \mu(1-\mu)^3(p_1 I_1 + p_4 I_4) + (1-\mu)^4 p_{14} I_{14} - cV_{14}.
\end{aligned} \tag{5}$$

In the first two equations, j belongs to a new index set Θ ($j \in \Theta$), where $\Theta = \{0, 1, 2, 3, 4, 13, 14\}$. From the calculations in Materials and Methods (Eqs. (14) and (15)), the pretreatment steady states of mutant strains are

$$\begin{aligned}
V_i &= \frac{1}{1-r_i}\frac{\mu}{1-\mu}V_0, \quad i = 1,2,3,4,\\
V_{13} &= \frac{1}{1-r_{13}}\left(\frac{\mu}{1-\mu}\right)^2\left(1 + \frac{r_1}{1-r_1} + \frac{r_3}{1-r_3}\right)V_0,\\
V_{14} &= \frac{1}{1-r_{14}}\left(\frac{\mu}{1-\mu}\right)^2\left(1 + \frac{r_1}{1-r_1} + \frac{r_4}{1-r_4}\right)V_0,
\end{aligned} \tag{6}$$

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

Using estimates of the relative fitness (assuming the fitness of wild-type virus is 1) of each mutant strain derived from a clinical study [32], 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, they only account for a very small fraction of the entire virus population. New

technologies such as pyrosequencing [63] may allow one to determine their frequency, but to our knowledge this has not been done yet.

Quasispecies Dynamics during Therapy

As suggested in previous sections, mutation during treatment does not have a strong effect on the evolution 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 alone cannot determine which strain(s) will dominate the virus population during therapy. Here we ignore all the mutations generated during therapy in Eq. (5). Then the multi-strain model under treatment becomes

$$\begin{aligned}
\frac{d}{dt}T(t) &= s - dT - \sum_i \beta_i V_i T, \\
\frac{d}{dt}I_i(t) &= \beta_i V_i T - \delta I_i, \\
\frac{d}{dt}V_i(t) &= (1-\epsilon_i)p_i I_i - cV_i,
\end{aligned} \tag{7}$$

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

The reproductive ratio for each strain under therapy is: $\mathcal{R}_i' = [(1-\epsilon_i)\beta_i p_i s]/(dc\delta), \ \ i \in \Theta$. 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, following the similar arguments in Supporting Information 1, it can be shown that only the viral strain with the largest reproductive ratio will persist during therapy if all the reproductive ratios are different from each other. All the other strains with lower reproductive ratios will die out. Mathematically, if $\mathcal{R}_i' > \max(1, \mathcal{R}_j')$, 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 greater than the reproductive ratios of other strains.

A few issues need 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 are generated by mutations. 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.

Discussion

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 current standard treatment. Specifically targeted antiviral therapy for HCV (STAT-C) has 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 used 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 interferon (IFN) antiviral response through its ability to block retinoic acid-inducible gene I and toll-like receptor-3 signaling [64, 65]. Therefore, the NS3-4A protease has become an ideal target for the development of new anti-HCV agents. Telaprevir (VX-950), a new protease inhibitor, has demonstrated substantial antiviral activity in clinical studies [16, 32–34]. Administration of telaprevir even in monotherapy resulted in $\sim$ 4-log reduction of the plasma viral load in HCV genotype 1 infected patients after 14 days [16]. However, drug resistant viral variants were detected within several days during the dosing period. The exact mechanisms underlying emergence of viral variants such a short time after initiation of therapy is still 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 STAT-C 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 the simplified two-strain model, we developed an analytical solution of the mutant frequency following treatment with telaprevir. We showed that the rapid increase of 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. With 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 significant impact on the evolution 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 substantially 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. 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.

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 involves 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.

The prevalence of the pre-existing viral variants may be an important factor that influences how quickly drug resistant viral strains emerge after drug administration. McPhee et al. [66] examined the baseline prevalence of pre-existing HCV variants resistant to protease inhibitors using a highly sensitive assay (limit of detection $< 0.1\%$ of the total population). In 3 of 8 patients, they detected the A156T variant at a frequency of 0.36%-0.75%. Cubero et al. [67] 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 by compensatory mutations (not considered in our models, see a recent paper [68] modeling 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 [69]. In the study of Cubero et al. [67], 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 pre-existence of mutant viral variants and the evolution of drug resistance during treatment requires more *in vitro* and *in vivo* studies.

From the calculation in (15), 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 mutations have seldom been identified so far. This raises the chance of success of an attractive 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) [70] and HIV treatment [71]. This idea has been recently confirmed in *in vitro* studies [72–74]. 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 [72], suggestive of a lack of cross resistance among the evaluated inhibitors. The data from a chimpanzee model of chronic HCV infection also support further investigation of combination therapy consisting of direct antiviral agents [75]. Therefore, combination of these specifically targeted antiviral drugs might be beneficial to HCV patients. More clinical data on toxicity and drug-drug-interactions are needed. 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 [76]. In clinical studies [34, 77], telaprevir was combined with PEG-IFN-alpha-2a and caused a continued antiviral response during the dosing period. 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 [34]. These results suggest that HCV variants with reduced sensitivity to telaprevir may remain sensitive to IFN plus RBV. Based on this, it seems that IFN and RBV will not be removed from antiviral regimens for HCV infection in the near future.

Materials and Methods

The Mutant Frequency after Drug Administration in the Two-Strain Model

Assuming T remains at a constant level, T_0, over a period of a few days after drug administration, and ignoring the term $\mu(1-\epsilon_s)p_sI_s$ in the V_r equation, Eq. (2) is reduced to the system

$$\begin{aligned}
\frac{d}{dt}I_s(t) &= \beta_s V_s T_0 - \delta I_s,\\
\frac{d}{dt}V_s(t) &= (1-\mu)(1-\epsilon_s)p_sI_s - cV_s,\\
\frac{d}{dt}I_r(t) &= \beta_r V_r T_0 - \delta I_r,\\
\frac{d}{dt}V_r(t) &= (1-\epsilon_r)p_rI_r - cV_r,
\end{aligned} \tag{8}$$

where $T_0 = c\delta/[(1-\mu)p_s\beta_s]$ is the pretreatment steady state of target cells in model (1). We note that the steady state T_0 does not rely on the drug resistant virus because backward mutation from drug resistant to wild-type was ignored in model (1).

The solution of the above system is

$$V_s(t) = C_1e^{\lambda_1 t} + C_2e^{\lambda_2 t}, \quad V_r(t) = C_3e^{\lambda_3 t} + C_4e^{\lambda_4 t},$$

where λ_i, $i = 1,2,3,4$, are the eigenvalues of the system, given by

$$\lambda_{1,2} = -\frac{c+\delta \pm \sqrt{\Delta_1}}{2}, \quad \lambda_{3,4} = -\frac{c+\delta \pm \sqrt{\Delta_2}}{2},$$

with

$$\Delta_1 = (c+\delta)^2 - 4[c\delta - (1-\epsilon_s)(1-\mu)p_s\beta_sT_0], \quad \Delta_2 = (c+\delta)^2 - 4[c\delta - (1-\epsilon_r)p_r\beta_rT_0],$$

which can be simplified to

$$\Delta_1 = (c+\delta)^2 - 4\epsilon_s c\delta, \quad \Delta_2 = (c+\delta)^2 - 4c\delta\left[1 - \frac{(1-\epsilon_r)\mathcal{R}_r}{(1-\mu)\mathcal{R}_s}\right].$$

In Supporting Information 2, we show that $(c+\delta)^2 > \Delta_2 > \Delta_1 > (c-\delta)^2 > 0$. Thus, $\lambda_i < 0$, $i = 1,2,3,4$.

The coefficients C_i, $i = 1,2,3,4$, are given by

$$C_{1,2} = \frac{\sqrt{\Delta_1} \mp \Big(c(1-2\epsilon_s)+\delta\Big)}{2\sqrt{\Delta_1}}V_s(0), \quad C_{3,4} = \frac{\sqrt{\Delta_2} \mp \Big(c(-1+\frac{2(1-\epsilon_r)}{1-\mu}\frac{\mathcal{R}_r}{\mathcal{R}_s})+\delta\Big)}{2\sqrt{\Delta_2}}V_r(0),$$

where

$$V_r(0) = \frac{\mu}{1-\mu-\frac{\mathcal{R}_r}{\mathcal{R}_s}}V_s(0).$$

We show in Supporting Information 2 that $C_i > 0$, $i = 1, 2, 3, 4$.

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

$$\Phi(t) = \frac{V_r(t)}{V_s(t) + V_r(t)} = \frac{C_3 e^{\lambda_3 t} + C_4 e^{\lambda_4 t}}{C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + C_3 e^{\lambda_3 t} + C_4 e^{\lambda_4 t}}, \tag{9}$$

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

Steady States of the Model with Hepatocyte Proliferation

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

$$\begin{aligned}
\frac{d}{dt}T(t) &= s + \rho_T T(1 - \frac{T + I_s + I_r}{T_{max}}) - dT - \beta_s V_s T - \beta_r V_r T, \\
\frac{d}{dt}I_s(t) &= \beta_s V_s T + \rho_s I_s(1 - \frac{T + I_s + I_r}{T_{max}}) - \delta I_s, \\
\frac{d}{dt}I_r(t) &= \beta_r V_r T + \rho_r I_r(1 - \frac{T + I_s + I_r}{T_{max}}) - \delta I_r, \\
\frac{d}{dt}V_s(t) &= (1 - \epsilon_s) p_s I_s - c V_s, \\
\frac{d}{dt}V_r(t) &= (1 - \epsilon_r) p_r I_r - c V_r.
\end{aligned} \tag{10}$$

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

$$\left[\frac{(1 - \epsilon_s)\beta_s p_s \bar{T}}{c} + \rho_s(1 - \frac{\bar{T} + \bar{I}_s + \bar{I}_r}{T_{max}}) - \delta\right]\bar{I}_s = 0 \tag{11}$$

and

$$\left[\frac{(1 - \epsilon_r)\beta_r p_r \bar{T}}{c} + \rho_r(1 - \frac{\bar{T} + \bar{I}_s + \bar{I}_r}{T_{max}}) - \delta\right]\bar{I}_r = 0. \tag{12}$$

If $\rho_s = \rho_r$, then it is obvious that the two strains cannot coexist because of $(1-\epsilon_s)\beta_s p_s < (1-\epsilon_r)\beta_r p_r$. If $\rho_s \neq \rho_r$, then it is possible that the two strains coexist. In this scenario, from (11) and (12) we have

$$\frac{(1 - \epsilon_s)\beta_s p_s \bar{T} - c\delta}{(1 - \epsilon_r)\beta_r p_r \bar{T} - c\delta} = \frac{\rho_s}{\rho_r},$$

which yields

$$\bar{T} = \frac{(\rho_s - \rho_r)c\delta}{(1 - \epsilon_r)\rho_s \beta_r p_r - (1 - \epsilon_s)\rho_r \beta_s p_s}.$$

Combining (11) or (12) with the T equation in Eq. (10), we can obtain the steady states $\bar{I}_s$ and $\bar{I}_r$. Since their expressions are complicated, we do not present them here.

The Multi-Strain Model and the Pretreatment Mutant Frequency

The full model with 4 possible mutation positions (without backward mutation) is

$$
\begin{aligned}
&\frac{d}{dt}T(t) = s - dT - \sum_j \beta_j V_j T,\\
&\frac{d}{dt}I_j(t) = \beta_j V_j T - \delta I_j,\\
&\frac{d}{dt}V_0(t) = (1-\mu)^4 p_0 I_0 - cV_0,\\
&\frac{d}{dt}V_i(t) = \mu(1-\mu)^3 p_0 I_0 + (1-\mu)^4 p_i I_i - cV_i, \quad i = 1,2,3,4,\\
&\frac{d}{dt}V_{ij}(t) = \mu^2(1-\mu)^2 p_0 I_0 + \mu(1-\mu)^3(p_i I_i + p_j I_j) + (1-\mu)^4 p_{ij} I_{ij} - cV_{ij},\\
&\qquad\qquad i,j = 1,2,3,4 \text{ and } i<j,\\
&\frac{d}{dt}V_{ijk}(t) = \mu^3(1-\mu) p_0 I_0 + \mu^2(1-\mu)^2(p_i I_i + p_j I_j + p_k I_k)\\
&\qquad +\mu(1-\mu)^3(p_{ij} I_{ij} + p_{ik} I_{ik} + p_{jk} I_{jk}) + (1-\mu)^4 p_{ijk} I_{ijk} - cV_{ijk},\\
&\qquad\qquad i,j,k = 1,2,3,4 \text{ and } i<j<k,\\
&\frac{d}{dt}V_{1234}(t) = \mu^4 p_0 I_0 + \mu^3(1-\mu)\sum_{i=1}^{4} p_i I_i + \mu^2(1-\mu)^2 \sum_{\substack{i,j=1,2,3,4\\ i<j}} p_{ij} I_{ij}\\
&\qquad +\mu(1-\mu)^3 \sum_{\substack{i,j,k=1,2,3,4\\ i<j<k}} p_{ijk} I_{ijk} + (1-\mu)^4 p_{1234} I_{1234} - cV_{1234}.
\end{aligned}
\tag{13}
$$

In the first two equations, the strain index j is in the set Ω ($j \in \Omega$), where Ω={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).

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 $r_i = \mathcal{R}_i/\mathcal{R}_0$, $i \in \Omega\backslash\{0\}$. 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 $[0,1]$.

A tedious but straightforward calculation yields the mutant frequency of the pre-existing viral

variants before treatment. The viral load of each strain is

$$
\begin{aligned}
V_i &= \frac{1}{1-r_i}\frac{\mu}{1-\mu}V_0, \quad i=1,2,3,4,\\
V_{ij} &= \frac{1}{1-r_{ij}}\left[(\frac{\mu}{1-\mu})^2V_0+\frac{\mu}{1-\mu}(r_iV_i+r_jV_j)\right], \quad i,j=1,2,3,4 \text{ and } i<j,\\
V_{ijk} &= \frac{1}{1-r_{ijk}}\left[(\frac{\mu}{1-\mu})^3V_0+(\frac{\mu}{1-\mu})^2(r_iV_i+r_jV_j+r_kV_k)\right.\\
&\quad \left.+\frac{\mu}{1-\mu}(r_{ij}V_{ij}+r_{ik}V_{ik}+r_{jk}V_{jk})\right], \quad i,j,k=1,2,3,4 \text{ and } i<j<k, \qquad (14)\\
V_{1234} &= \frac{1}{1-r_{1234}}\left[(\frac{\mu}{1-\mu})^4V_0+(\frac{\mu}{1-\mu})^3\sum_{i=1}^{4}r_iV_i+(\frac{\mu}{1-\mu})^2\sum_{\substack{i,j=1,2,3,4\\ i<j}}r_{ij}V_{ij}\right.\\
&\quad \left.+\frac{\mu}{1-\mu}\sum_{\substack{i,j,k=1,2,3,4\\ i<j<k}}r_{ijk}V_{ijk}\right],
\end{aligned}
$$

where V_0 is the steady sate 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}):

$$
\begin{aligned}
V_{ij} &= \frac{1}{1-r_{ij}}(\frac{\mu}{1-\mu})^2(1+\sigma_i+\sigma_j)V_0, \quad i,j=1,2,3,4 \text{ and } i<j,\\
V_{ijk} &= \frac{1}{1-r_{ijk}}(\frac{\mu}{1-\mu})^3\left[1+\sum_{l=i,j,k}\sigma_l+\sum_{\substack{l,m=i,j,k\\ l<m}}\sigma_{lm}(1+\sigma_l+\sigma_m)\right]V_0,\\
&\qquad\qquad i,j,k=1,2,3,4 \text{ and } i<j<k, \qquad (15)\\
V_{1234} &= \frac{1}{1-r_{1234}}(\frac{\mu}{1-\mu})^4\left[1+\sum_{i=1}^{4}\sigma_i+\sum_{\substack{i,j=1,2,3,4\\ i<j}}\sigma_{ij}(1+\sigma_i+\sigma_j)\right.\\
&\quad \left.+\sum_{\substack{i,j,k=1,2,3,4\\ i<j<k}}\sigma_{ijk}\Big(1+\sum_{l=i,j,k}\sigma_l+\sum_{\substack{l,m=i,j,k\\ l<m}}\sigma_{lm}(1+\sigma_l+\sigma_m)\Big)\right]V_0,
\end{aligned}
$$

where $\sigma_i = r_i/(1-r_i)$, $i \in \Omega\backslash\{0\}$. It follows that the pretreatment steady state level of an m-mutant viral variant is of the order of μ^m. The ratio of the m-mutant variant to wild-type virus, $V_{m-mutant}/V_0$, 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 depends 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 $\Phi_i = V_i/V_{total}$, where $V_{total} = \sum_{i\in\Omega} V_i$. The frequency Φ_i depends on the mutation rate μ and the relative fitness of all mutant strains, r_i, $i \in \Omega\backslash\{0\}$.

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, $V_{m-mutant}$, has mutations occurring at the first m positions. Then the pre-existing steady state of this strain is

$$V_{m-mutant} = \frac{1}{1 - r_{m-mutant}} \Big[(\frac{\mu}{1-\mu})^m V_0 + (\frac{\mu}{1-\mu})^{m-1} \sum_{i=1}^{m} r_i V_i + (\frac{\mu}{1-\mu})^{m-2} \sum_{\substack{i,j=1,\ldots,m \\ i<j}} r_{ij} V_{ij}$$

$$+ (\frac{\mu}{1-\mu})^{m-3} \sum_{\substack{i,j,k=1,\ldots,m \\ i<j<k}} r_{ijk} V_{ijk} + \ldots \Big].$$

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 (14). In this way, we can calculate the frequency of all the pre-existing variants in a general model with n possible mutations.

Acknowledgements Portions of this work were performed under the auspices of the U.S. Department of Energy under contract DE-AC52-06NA25396. This work was supported by NIH grants RR06555, AI28433 (A.S.P) and P20-RR18754 (R.M.R).

References

[1] World Health Organization. Hepatitis C. Fact sheet No. 164. Revised October 2000. http://www.who.int/mediacentre/factsheets/fs164/en/index.html .

[2] World Health Organization (2000) Hepatitis C - Global prevalence (update). Weekly Epidemiol. Rec. 75:18-19 .

[3] Alter HJ (2005) HCV natural history: the retrospective and prospective in perspective. J Hepatol 43:550–2.

[4] Foster GR (2004) Past, present, and future hepatitis C treatments. Semin Liver Dis 24 Suppl 2:97–104.

[5] Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, et al. (2002) Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 347:975–82.

[6] Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, et al. (2001) Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 358:958–65.

[7] Herrmann E, Lee J, Marinos G, Modi M, Zeuzem S (2003) Effect of ribavirin on hepatitis C viral kinetics in patients treated with pegylated interferon. Hepatology 37:1351–8.

[8] Parfieniuk A, Jaroszewicz J, Flisiak R (2007) Specifically targeted antiviral therapy for hepatitis C virus. World J Gastroenterol 13:5673–81.

[9] De Francesco R, Migliaccio G (2005) Challenges and successes in developing new therapies for hepatitis C. Nature 436:953–60.

[10] Lamarre D, Anderson PC, Bailey M, Beaulieu P, Bolger G, et al. (2003) An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. Nature 426:186–9.

[11] Hinrichsen H, Benhamou Y, Wedemeyer H, Reiser M, Sentjens RE, et al. (2004) Short-term antiviral efficacy of BILN 2061, a hepatitis C virus serine protease inhibitor, in hepatitis C genotype 1 patients. Gastroenterology 127:1347–55.

[12] Reiser M, Hinrichsen H, Benhamou Y, Reesink HW, Wedemeyer H, et al. (2005) Antiviral efficacy of NS3-serine protease inhibitor BILN-2061 in patients with chronic genotype 2 and 3 hepatitis C. Hepatology 41:832–5.

[13] Sarrazin C, Rouzier R, Wagner F, Forestier N, Larrey D, et al. (2007) SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon alpha-2b for genotype 1 nonresponders. Gastroenterology 132:1270–8.

[14] Lin K, Perni RB, Kwong AD, Lin C (2006) VX-950, a novel hepatitis C virus (HCV) NS3-4A protease inhibitor, exhibits potent antiviral activities in HCV replicon cells. Antimicrob Agents Chemother 50:1813–22.

[15] Perni RB, Almquist SJ, Byrn RA, Chandorkar G, Chaturvedi PR, et al. (2006) Preclinical profile of VX-950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3-4A serine protease. Antimicrob Agents Chemother 50:899–909.

[16] Reesink HW, Zeuzem S, Weegink CJ, Forestier N, van Vliet A, et al. (2006) Rapid decline of viral RNA in hepatitis C patients treated with VX-950: a phase Ib, placebo-controlled, randomized study. Gastroenterology 131:997–1002.

[17] Pawlotsky J (2008) Hepatitis virus resistance. In: Fong I, Drlica K, editors, Antimicrobial Resistance and Implications for the 21st Century, Springer. pp. 291–323.

[18] Farci P, Strazzera R, Alter HJ, Farci S, Degioannis D, et al. (2002) Early changes in hepatitis C viral quasispecies during interferon therapy predict the therapeutic outcome. Proc Natl Acad Sci U S A 99:3081–6.

[19] Martell M, Esteban JI, Quer J, Genesca J, Weiner A, et al. (1992) Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. J Virol 66:3225–9.

[20] Pawlotsky JM (2006) Hepatitis C virus population dynamics during infection. Curr Top Microbiol Immunol 299:261–84.

[21] Simmonds P (2004) Genetic diversity and evolution of hepatitis C virus–15 years on. J Gen Virol 85:3173–88.

[22] Domingo E, Holland J (1994) Mutation rates and rapid evolution of RNA viruses. In: Morse S, editor, Evolutionary Biology of Viruses, New York: Raven Press. pp. 161–184.

[23] Holland JJ, De La Torre JC, Steinhauer DA (1992) RNA virus populations as quasispecies. Curr Top Microbiol Immunol 176:1–20.

[24] Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439:344–8.

[25] Hirsch MS, Conway B, D'Aquila RT, Johnson VA, Brun-Vezinet F, et al. (1998) Antiretroviral drug resistance testing in adults with HIV infection: implications for clinical management. International AIDS Society–USA Panel. JAMA 279:1984–91.

[26] Rong L, Feng Z, Perelson AS (2007) Emergence of HIV-1 drug resistance during antiretroviral treatment. Bull Math Biol 69:2027–60.

[27] Lin C, Gates CA, Rao BG, Brennan DL, Fulghum JR, et al. (2005) In vitro studies of cross-resistance mutations against two hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061. J Biol Chem 280:36784–91.

[28] Lin C, Lin K, Luong YP, Rao BG, Wei YY, et al. (2004) In vitro resistance studies of hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061: structural analysis indicates different resistance mechanisms. J Biol Chem 279:17508–14.

[29] Lu L, Pilot-Matias TJ, Stewart KD, Randolph JT, Pithawalla R, et al. (2004) Mutations conferring resistance to a potent hepatitis C virus serine protease inhibitor in vitro. Antimicrob Agents Chemother 48:2260–6.

[30] Malcolm BA, Liu R, Lahser F, Agrawal S, Belanger B, et al. (2006) SCH 503034, a mechanism-based inhibitor of hepatitis C virus NS3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells. Antimicrob Agents Chemother 50:1013–20.

[31] Tong X, Chase R, Skelton A, Chen T, Wright-Minogue J, et al. (2006) Identification and analysis of fitness of resistance mutations against the HCV protease inhibitor SCH 503034. Antiviral Res 70:28–38.

[32] Sarrazin C, Kieffer TL, Bartels D, Hanzelka B, Muh U, et al. (2007) Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. Gastroenterology 132:1767–77.

[33] Forestier N, Reesink HW, Weegink CJ, McNair L, Kieffer TL, et al. (2007) Antiviral activity of telaprevir (VX-950) and peginterferon alfa-2a in patients with hepatitis C. Hepatology 46:640–8.

[34] Kieffer TL, Sarrazin C, Miller JS, Welker MW, Forestier N, et al. (2007) Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis C virus replication in patients. Hepatology 46:631–9.

[35] Ribeiro RM, Bonhoeffer S, Nowak MA (1998) The frequency of resistant mutant virus before antiviral therapy. AIDS 12:461–5.

[36] Bonhoeffer S, Nowak MA (1997) Pre-existence and emergence of drug resistance in HIV-1 infection. Proc Biol Sci 264:631–7.

[37] Nowak MA, Bonhoeffer S, Shaw GM, May RM (1997) Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations. J Theor Biol 184:203–17.

[38] Ribeiro RM, Bonhoeffer S (2000) Production of resistant HIV mutants during antiretroviral therapy. Proc Natl Acad Sci U S A 97:7681–6.

[39] McLean AR, Nowak MA (1992) Competition between zidovudine-sensitive and zidovudine-resistant strains of HIV. AIDS 6:71–9.

[40] Rong L, Gilchrist MA, Feng Z, Perelson AS (2007) Modeling within-host HIV-1 dynamics and the evolution of drug resistance: trade-offs between viral enzyme function and drug susceptibility. J Theor Biol 247:804–18.

[41] Wahl LM, Nowak MA (2000) Adherence and drug resistance: predictions for therapy outcome. Proc Biol Sci 267:835–43.

[42] Murray J, Perelson A (2005) Human immunodeficiency virus: quasi-species and drug resistance. Multiscale Model Simul 3:300–311.

[43] Nowak MA, May RM, Anderson RM (1990) The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease. AIDS 4:1095–103.

[44] Lindenbach BD, Rice CM (2005) Unravelling hepatitis C virus replication from genome to function. Nature 436:933–8.

[45] Pawlotsky JM, Chevaliez S, McHutchison JG (2007) The hepatitis C virus life cycle as a target for new antiviral therapies. Gastroenterology 132:1979–98.

[46] Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, et al. (1998) Binding of hepatitis C virus to CD81. Science 282:938–41.

[47] Scarselli E, Ansuini H, Cerino R, Roccasecca RM, Acali S, et al. (2002) The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. Embo J 21:5017–25.

[48] Egger D, Wolk B, Gosert R, Bianchi L, Blum HE, et al. (2002) Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. J Virol 76:5974–84.

[49] Bartenschlager R, Frese M, Pietschmann T (2004) Novel insights into hepatitis C virus replication and persistence. Adv Virus Res 63:71–180.

[50] Domingo E (1996) Biological significance of viral quasispecies. Viral Hepatitis Rev 2:247–261.

[51] Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, et al. (1998) Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science 282:103–7.

[52] Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD (1996) HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271:1582–6.

[53] Anderson R, May R (1992) Infectious Diseases of Humans: Dynamics and Control. USA: Oxford University Press.

[54] Clavel F, Race E, Mammano F (2000) HIV drug resistance and viral fitness. Adv Pharmacol 49:41–66.

[55] Gillespie J (1998) Population Genetics: A Concise Guide. Johns Hopkins University Press.

[56] Chu H, Herrmann E, Reesink H, Forestier N, Weegink C, et al. (2005) Pharmacokinetics of VX-950, and its effect on hepatitis C viral dynamics. Hepatology 42 Suppl 1:S694.

[57] Gabrielson J, Weiner D (2000) Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications. Stockholm: Swedish Pharmaceutical Press.

[58] Shen L, Peterson S, Sedaghat AR, McMahon MA, Callender M, et al. (2008) Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. Nat Med 14:762–6.

[59] 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. In: Digestive Disease Week 2007. Washington DC, May 19-24, 2007.

[60] Michalopoulos GK, DeFrances MC (1997) Liver regeneration. Science 276:60–6.

[61] Dahari H, Lo A, Ribeiro RM, Perelson AS (2007) Modeling hepatitis C virus dynamics: liver regeneration and critical drug efficacy. J Theor Biol 247:371–81.

[62] Dahari H, Ribeiro RM, Perelson AS (2007) Triphasic decline of hepatitis C virus RNA during antiviral therapy. Hepatology 46:16–21.

[63] Elahi E, Ronaghi M (2004) Pyrosequencing: a tool for DNA sequencing analysis. Methods Mol Biol 255:211–9.

[64] Foy E, Li K, Wang C, Sumpter J R, Ikeda M, et al. (2003) Regulation of interferon regulatory factor-3 by the hepatitis C virus serine protease. Science 300:1145–8.

[65] Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, et al. (2005) Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature 437:1167–72.

[66] McPhee F, Hernandez D, Zhai G, Friborg J, Yu F, et al. Pre-existence of substitutions conferring resistance to HCV NS3 protease inhibitors. In: 1st International Workshop on Hepatitis C Resistance and New Compounds. Boston, MA, October 25-26, 2006.

[67] Cubero M, Esteban JI, Otero T, Sauleda S, Bes M, et al. (2008) Naturally occurring NS3-protease-inhibitor resistant mutant A156T in the liver of an untreated chronic hepatitis C patient. Virology 370:237–45.

[68] Handel A, Regoes RR, Antia R (2006) The role of compensatory mutations in the emergence of drug resistance. PLoS Comput Biol 2:e137.

[69] Yi M, Tong X, Skelton A, Chase R, Chen T, et al. (2006) 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 281:8205–15.

[70] Keeffe EB, Dieterich DT, Pawlotsky JM, Benhamou Y (2008) Chronic hepatitis B: preventing, detecting, and managing viral resistance. Clin Gastroenterol Hepatol 6:268–74.

[71] Yeni PG, Hammer SM, Carpenter CC, Cooper DA, Fischl MA, et al. (2002) Antiretroviral treatment for adult HIV infection in 2002: updated recommendations of the International AIDS Society-USA Panel. JAMA 288:222–35.

[72] McCown MF, Rajyaguru S, Le Pogam S, Ali S, Jiang WR, et al. (2008) The hepatitis C virus replicon presents a higher barrier to resistance to nucleoside analogs than to nonnucleoside polymerase or protease inhibitors. Antimicrob Agents Chemother 52:1604–12.

[73] Wyles DL, Kaihara KA, Schooley RT (2008) Synergy of a hepatitis C virus (HCV) NS4A antagonist in combination with HCV protease and polymerase inhibitors. Antimicrob Agents Chemother 52:1862–4.

[74] Wyles DL, Kaihara KA, Vaida F, Schooley RT (2007) Synergy of small molecular inhibitors of hepatitis C virus replication directed at multiple viral targets. J Virol 81:3005–8.

[75] Olsen D, Carroll S, Handt L, Ludmerer S, Graham D, et al. HCV Antiviral Activity and Resistance Analysis in Chronically infected Chimpanzees treated with NS3/4A Protease and NS5B Polymerase Inhibitors. In: 42nd Annual Meeting of the European Association for the Study of the Liver (EASL). Barcelona, Spain, April 11-15, 2007.

[76] Lin K, Kwong AD, Lin C (2004) 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 48:4784–92.

[77] Lawitz E, Rodriguez-Torres M, Muir AJ, Kieffer TL, McNair L, et al. (2008) Antiviral effects and safety of telaprevir, peginterferon alfa-2a, and ribavirin for 28 days in hepatitis C patients. J Hepatol .

Table 1. Effects of backward mutation on the steady states and the mutant frequency before treatment

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

Note: Model parameters used to obtain the steady states of V_s and V_r are: $d = 0.01$ day^{-1}, $p_s = 10$ virions cell^{-1} day^{-1}, $\mu = 10^{-4}$, $c = 6.2$ day^{-1}, $\delta = 0.24$ 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 $\beta_s = 10^{-7}$ mL day^{-1} virions^{-1} and $s = 7.5 \times 10^5$ cells mL^{-1} day^{-1}. For simplicity, we assume that wild-type and resistant viruses differ only in their replication capacities. We choose $\beta_r = \beta_s = 10^{-7}$ mL day^{-1} virions^{-1} and $p_r = 6$ virions cell^{-1} 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 [32]).

Table 2. 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	2.85×10^{-6}
36/156	0.67	1.53×10^{-6}

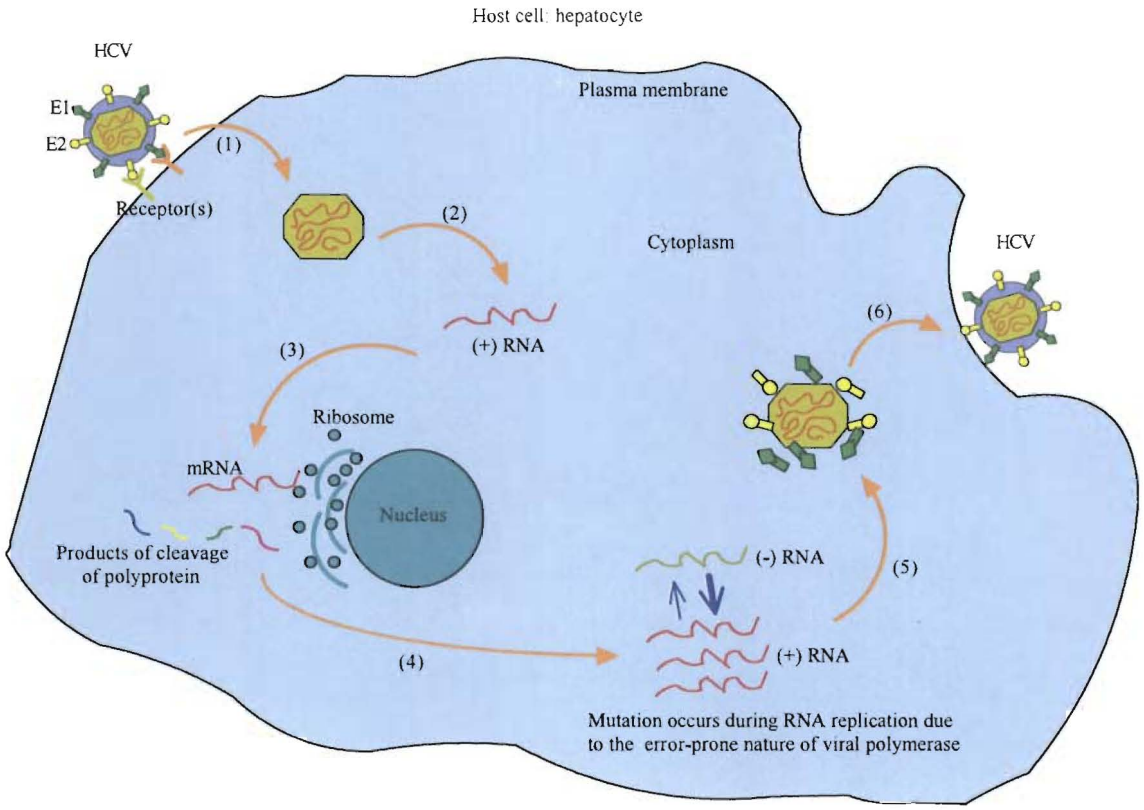

Figure 1: HCV life cycle. (1) Following viral binding, receptor-mediated endocytosis and membrane fusion, nucleocapsid enters into the cytoplasm of the host cell; (2) Uncoating of nucleocapsid exposes a positive-strand RNA genome; (3) Internal ribosomal entry site (IRES)-mediated translation of the viral 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.

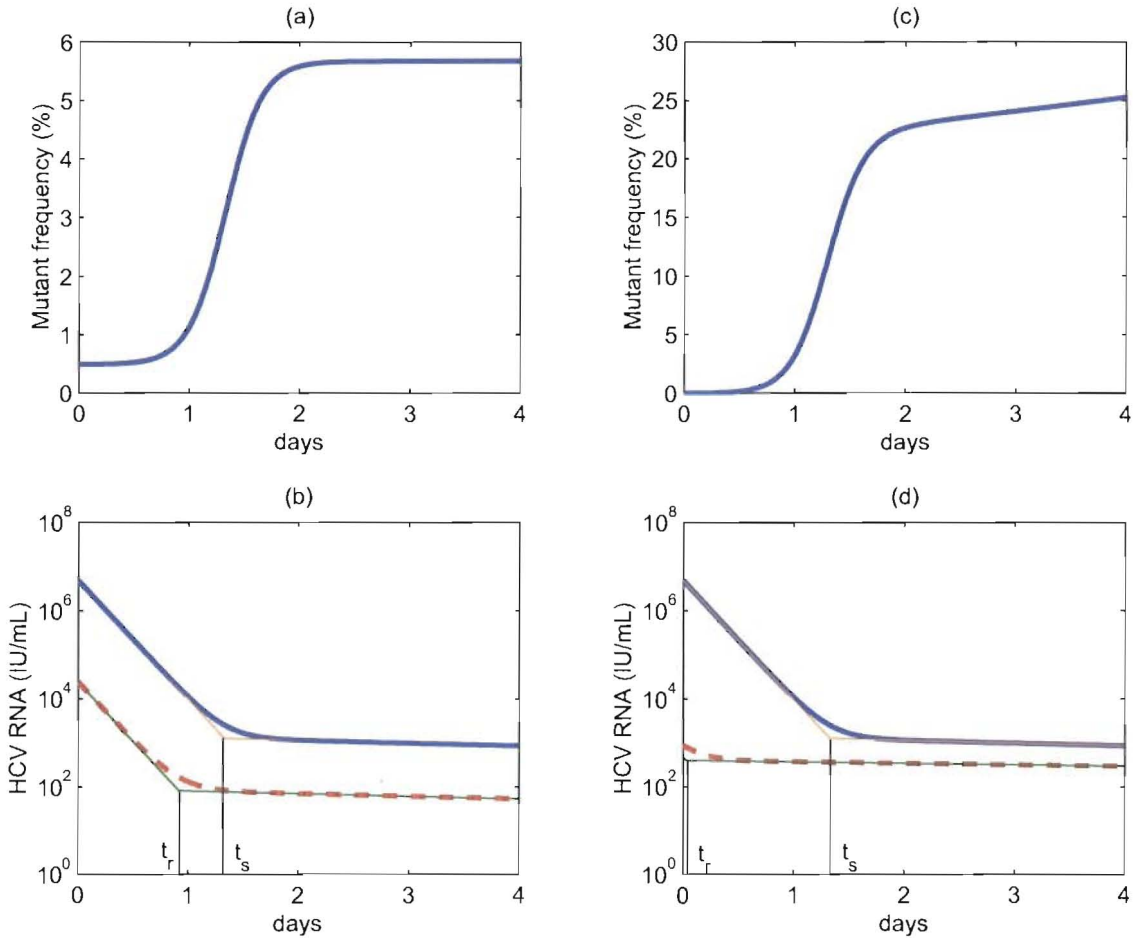

Figure 2: Increase of the mutant frequency (upper panels; see Eq. (9)) and changes of the viral levels (lower panels; blue solid line is wild-type virus, red dashed line is mutant virus) following treatment with telaprevir (assuming the target cell level is constant, see Eq. (8)). t_s is the time at 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: $c = 6.2$ day^{-1}, $\delta = 0.14$ day^{-1}, $\mu = 10^{-4}$ per copied nucleotide, $\epsilon_s = 0.9997$, and the Hill coefficient is $h = 2$. Left column: the mutant, for example V36A/M, confers 3.5-fold resistance and $\mathcal{R}_r/\mathcal{R}_s$=0.98 [32]. We obtained the eigenvalues $\lambda_1 = -6.2$, $\lambda_2 = -0.14$, $\lambda_3 = -6.2005$, $\lambda_4 = -0.1395$, and $t_s = 1.33$ day, $t_r = 0.92$ day. Right column: the mutant, for example A156V/T, confers 466-fold resistance and $\mathcal{R}_r/\mathcal{R}_s$=0.45 [32]. We obtained the eigenvalues $\lambda_1 = -6.2$, $\lambda_2 = -0.14$, $\lambda_3 = -6.2628$, $\lambda_4 = -0.0772$, and $t_s = 1.33$ day, $t_r = 0.03$ day. The increase of the mutant frequency following therapy is due to a longer first-phase viral decline of wild-type virus.

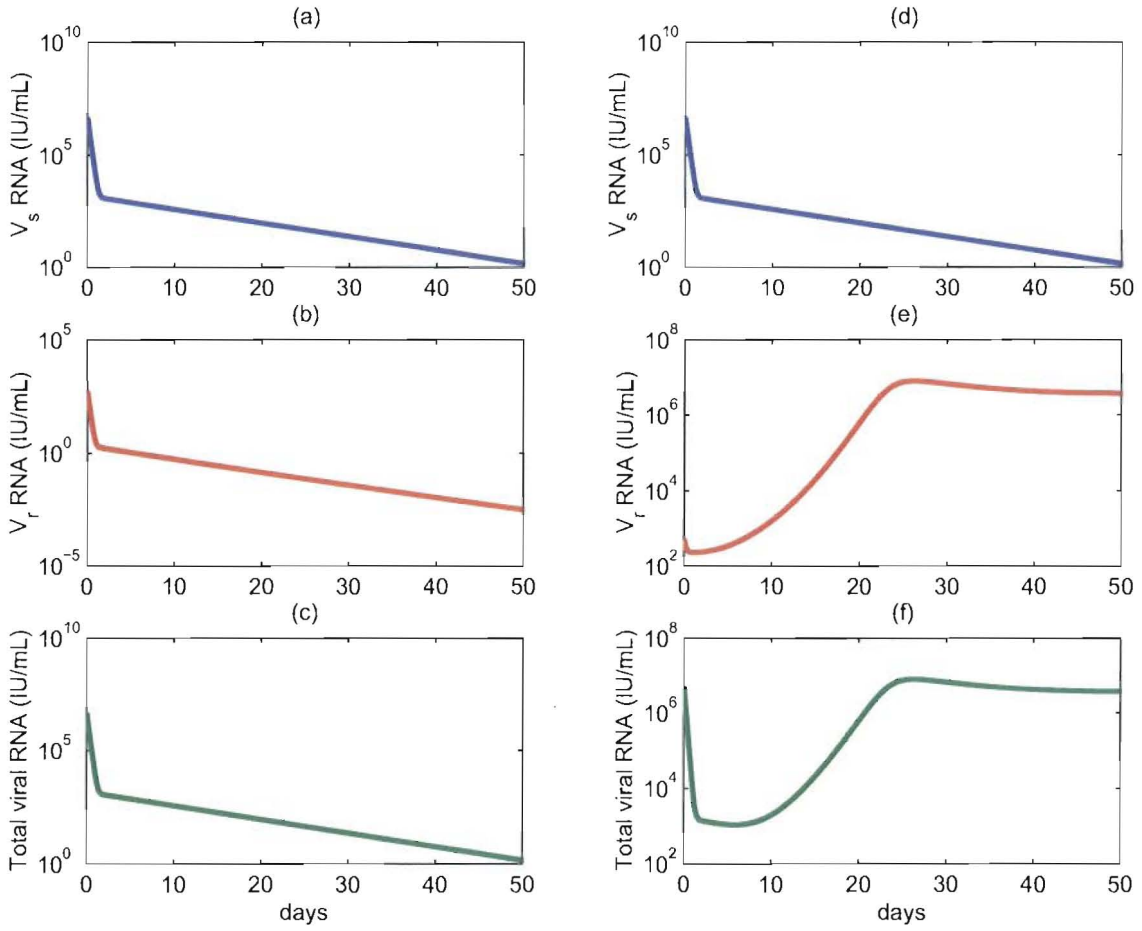

Figure 3: Competition between wild-type (upper panels) and drug resistant virus (middle panels) during therapy. The total viral levels are plotted in the lower panels. Left column: assuming the mutant, for example V36A/M, confers 3.5-fold resistance and $\mathcal{R}_r/\mathcal{R}_s$=0.98. Both wild-type 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. 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: $s = 7.5 \times 10^5$ cells mL^{-1} day^{-1}, $d = 0.01$ day^{-1}, $\beta_s = \beta_r = 10^{-7}$ mL day^{-1} virions^{-1}, $\mu = 10^{-4}$ per copied nucleotide, $c = 6.2$ day^{-1}, $\delta = 0.14$ day^{-1}, $p_s = 10$ virions cell^{-1} day^{-1}, $\epsilon_s = 0.9997$, and the Hill coefficient is $h = 2$.

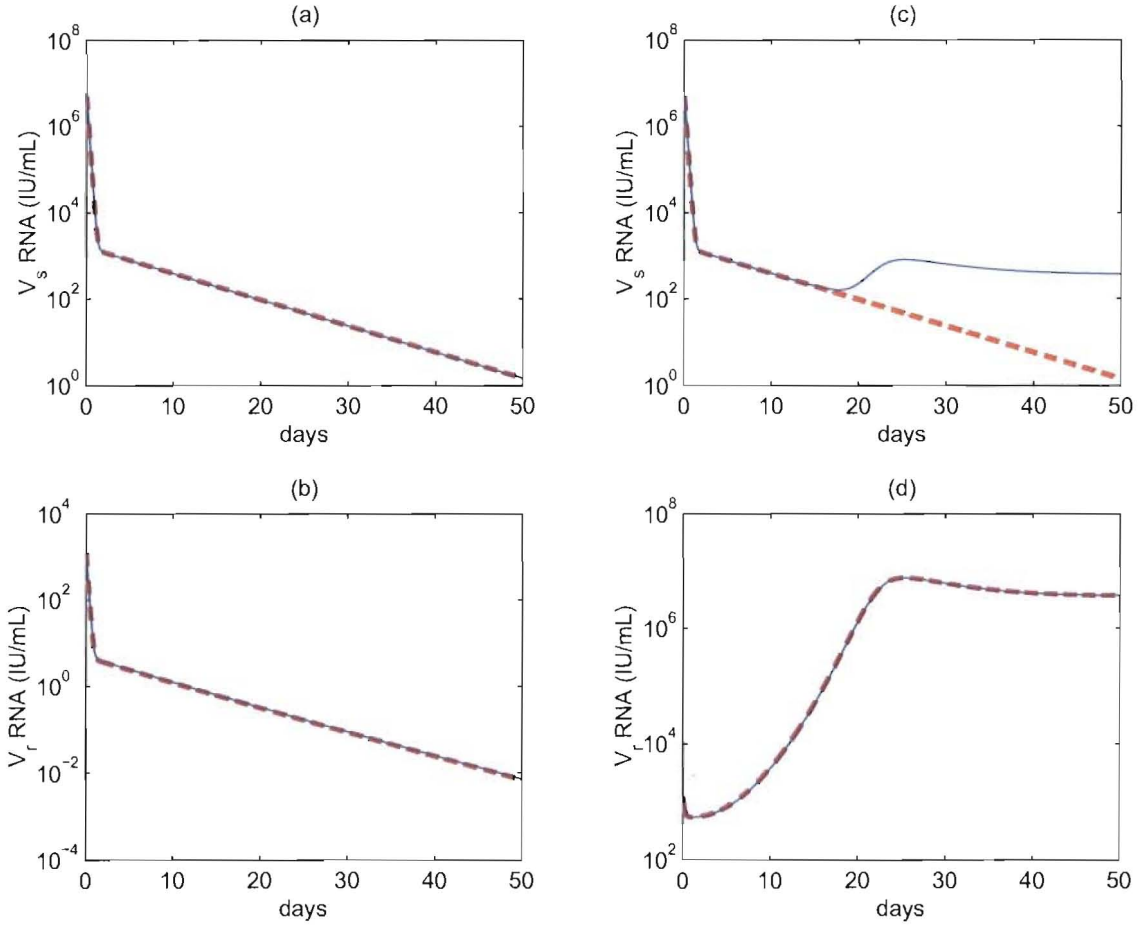

Figure 4: Contribution of mutation to the evolution of wild-type and drug resistant virus during treatment: assuming there is no mutation during treatment (thick dashed line) and there exist both forward and backward mutations during therapy (thin solid line). Left column: assuming the mutant confers 3.5-fold resistance and $\mathcal{R}_r/\mathcal{R}_s$=0.98. The solid and the dashed lines almost lap over, 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. Mutation still does not contribute largely 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: $s = 7.5 \times 10^5$ cells mL^{-1} day^{-1}, $d = 0.01$ day^{-1}, $\beta_s = \beta_r = 10^{-7}$ mL day^{-1} virions^{-1}, $\mu = 10^{-4}$ per copied nucleotide, $c = 6.2$ day^{-1}, $\delta = 0.14$ day^{-1}, $p_s = 10$ virions cell^{-1} day^{-1}, $\epsilon_s = 0.9997$, and $h = 2$.

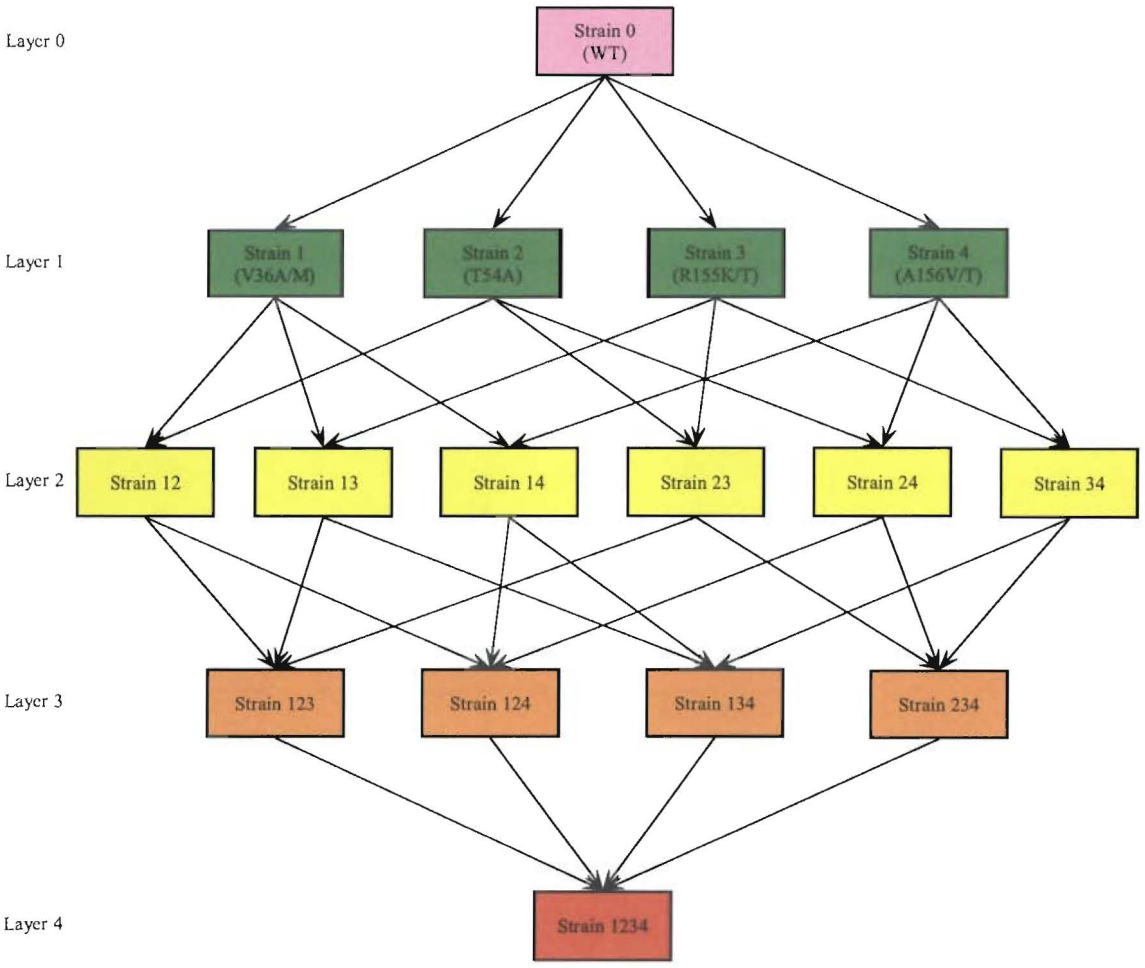

Figure 5: Mutation diagram: from the topmost (layer 0) to the bottommost (layer 4) are the wild-type, single-mutant, double-mutant, 3-mutant, and 4-mutant strains. Backward mutation is not considered. Mutations between strains that lead to transitions beyond one layer are not shown but are considered 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.

Supporting Information 1: Steady State Analysis of the Two-Strain Model

There are three possible steady states of the two-strain model before treatment: the infection-free steady state

$$E_0 = (\bar{T}, \bar{I}_s, \bar{V}_s, \bar{I}_r, \bar{V}_r) = (\frac{s}{d}, 0, 0, 0, 0),$$

the boundary steady state (only the drug-resistant strain is present)

$$E_r = \left(\frac{c\delta}{\beta_r p_r}, 0, 0, (\mathcal{R}_r - 1)\frac{dc}{\beta_r p_r}, (\mathcal{R}_r - 1)\frac{d}{\beta_r}\right),$$

and the interior steady state (coexistence of both wild-type and drug-resistant strains)

$$E_c = \left(\tilde{T}, \frac{\beta_s}{\delta}\tilde{T}\tilde{V}_s, \tilde{V}_s, \frac{\beta_r}{\delta}\tilde{T}\tilde{V}_r, \tilde{V}_r\right),$$

where

$$\tilde{T} = \frac{c\delta}{(1-\mu)\beta_s p_s},$$
$$\tilde{V}_s = \frac{[(1-\mu)\mathcal{R}_s - 1]d}{\beta_s + \beta_r \frac{\mu}{1-\mu-r}},$$
$$\tilde{V}_r = \frac{\mu}{1-\mu-r}\tilde{V}_s.$$

Here $r = \mathcal{R}_r/\mathcal{R}_s$, where $\mathcal{R}_s = \frac{s\beta_s p_s}{dc\delta}$ and $\mathcal{R}_r = \frac{s\beta_r p_r}{dc\delta}$ are the basic reproductive ratios of the wild-type strain and the drug-resistant strain, respectively. Thus, r represents the relative fitness of drug resistant to wild-type virus.

It is clear that E_r exists if and only if $\mathcal{R}_r > 1$, and E_c exists if and only if $\mathcal{R}_s > \frac{1}{1-\mu}$ and $r < 1-\mu$. We will show below that these existence conditions also provide threshold conditions for the stability of the steady states.

Linearizing the two-strain model about the steady state $\bar{E}$, we get

$$\frac{d}{dt}T(t) = -(d + \beta_s\bar{V}_s + \beta_r\bar{V}_r)T - \beta_s\bar{T}V_s - \beta_r\bar{T}V_r,$$
$$\frac{d}{dt}I_s(t) = \beta_s\bar{V}_s T - \delta I_s + \beta_s\bar{T}V_s,$$
$$\frac{d}{dt}I_r(t) = \beta_r\bar{V}_r T - \delta I_r + \beta_r\bar{T}V_r,$$
$$\frac{d}{dt}V_s(t) = (1-\mu)p_s I_s - cV_s,$$
$$\frac{d}{dt}V_r(t) = \mu p_s I_s + p_r I_r - cV_r.$$

The corresponding characteristic equation is

$$\begin{vmatrix} -(d+\beta_s\bar{V}_s+\beta_r\bar{V}_r)-\lambda & 0 & 0 & -\beta_s\bar{T} & -\beta_r\bar{T} \\ \beta_s\bar{V}_s & -\delta-\lambda & 0 & \beta_s\bar{T} & 0 \\ \beta_r\bar{V}_r & 0 & -\delta-\lambda & 0 & \beta_r\bar{T} \\ 0 & (1-\mu)p_s & 0 & -c-\lambda & 0 \\ 0 & \mu p_s & p_r & 0 & -c-\lambda \end{vmatrix} = 0, \qquad \text{(SI1-1)}$$

where λ denotes the eigenvalue of the linearized system.

(i) The infection-free steady state E_0 is locally asymptotically stable if $\mathcal{R}_s < 1/(1-u)$ and $\mathcal{R}_r < 1$, and it is unstable if $\mathcal{R}_s > 1/(1-u)$ or $\mathcal{R}_r > 1$.

Evaluating the characteristic equation (SI1-1) at $\bar{E} = E_0$, we get

$$(\lambda+d)\Big[(\lambda+c)(\lambda+\delta)-(1-\mu)\mathcal{R}_s c\delta\Big]\Big[(\lambda+c)(\lambda+\delta)-\mathcal{R}_r c\delta\Big] = 0. \qquad \text{(SI1-2)}$$

Equation (SI1-2) has one negative solution $-d$, and all other solutions are determined by either $(\lambda+c)(\lambda+\delta) = (1-\mu)\mathcal{R}_s c\delta$ or $(\lambda+c)(\lambda+\delta) = \mathcal{R}_r c\delta$. If $\mathcal{R}_s < 1/(1-u)$ and $\mathcal{R}_r < 1$, then all the solutions have negative real parts by comparing the moduli of both sides of the equations. This shows that E_0 is stable under the given conditions.

When one of the inequalities is reversed, e.g., $\mathcal{R}_s > 1/(1-\mu)$, we define an auxiliary function $f(\lambda) = (\lambda+c)(\lambda+\delta) - (1-\mu)\mathcal{R}_s c\delta$. Then $f(\lambda)$ has at least one positive root since $f(0) < 0$ and $f(\lambda) \to +\infty$ as $\lambda \to +\infty$. It follows that the equation $(\lambda+c)(\lambda+\delta) = (1-\mu)\mathcal{R}_s c\delta$ has at least one positive solution. Therefore, E_0 is unstable.

(ii) The steady state in which only drug resistant virus is present, E_r, exists if and only if $\mathcal{R}_r > 1$. It is locally asymptotically stable if $r > 1-\mu$ and unstable if $r < 1-\mu$.

We substitute the steady state E_r into the characteristic equation (SI1-1) and simplify it to the following equation:

$$\Big[(\lambda+c)(\lambda+\delta) - \frac{1-\mu}{r}c\delta\Big]\Big[(\lambda+d\mathcal{R}_r)(\lambda+c)(\lambda+\delta) - (\lambda+d)c\delta\Big] = 0. \qquad \text{(SI1-3)}$$

The eigenvalues are determined by the equation

$$(\lambda+c)(\lambda+\delta) = \frac{1-\mu}{r}c\delta, \qquad \text{(SI1-4)}$$

or

$$(\lambda+d\mathcal{R}_r)(\lambda+c)(\lambda+\delta) = (\lambda+d)c\delta. \qquad \text{(SI1-5)}$$

By comparing the moduli of both sides of the equation (SI1-4) we show that all the solutions have negative real parts if $r > 1-\mu$, i.e., $\mathcal{R}_r > (1-\mu)\mathcal{R}_s$. Similarly, all the solutions of the

equation (SI1-5) have negative real parts whenever E_r exists, i.e., $\mathcal{R}_r > 1$. Thus, all the solutions of the characteristic equation (SI1-3) have negative real parts when $\mathcal{R}_r > 1$ and $\mathcal{R}_r > (1-\mu)\mathcal{R}_s$. Therefore, the steady state E_r is locally asymptotically stable if $r > 1-\mu$. By the same arguments as used in (i), the equation (SI1-4) has at least one positive solution when $r < 1-\mu$. This implies that E_r is unstable when $r < 1-\mu$.

(iii) The coexistence steady state E_c exists and is locally asymptotically stable if and only if $\mathcal{R}_s > 1/(1-\mu)$ and $r < 1-\mu$.

Evaluating the determinant in (SI1-1) at $\bar{E} = E_c$, we obtain the simplified characteristic equation after tedious calculations:

$$\left[(\lambda+c)(\lambda+\delta) - \frac{r}{1-\mu}c\delta\right]\left[\left(\lambda + d(1-\mu)\mathcal{R}_s\right)(\lambda+c)(\lambda+\delta) - (\lambda+d)c\delta\right] = 0.$$

Therefore, the solutions are determined by the equation

$$(\lambda+c)(\lambda+\delta) = \frac{r}{1-\mu}c\delta, \tag{SI1-6}$$

or

$$\left[\lambda + d(1-\mu)\mathcal{R}_s\right](\lambda+c)(\lambda+\delta) = (\lambda+d)c\delta. \tag{SI1-7}$$

It follows from the assumption $r < 1-\mu$ that all the solutions of (SI1-6) have negative real parts. Considering the moduli of both sides of (SI1-7), λ must fall in the left half plane because $\mathcal{R}_s > 1/(1-\mu)$. Therefore, all the solutions of the characteristic equation are in the left half plane, and hence E_c is locally asymptotically stable.

On the other hand, if either of the two conditions, $\mathcal{R}_s > 1/(1-\mu)$ and $r < 1-\mu$, does not hold, then E_c does not exist. This shows that the steady state E_c is locally asymptotically stable whenever it exists.

Supporting Information 2: Several Inequalities Used in the Analysis of the Two-Strain Model

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

From $\Delta_1 = (c+\delta)^2 - 4\epsilon_s c\delta$, it is clear that $(c+\delta)^2 > \Delta_1 > (c-\delta)^2$. Thus, we have $\lambda_1 = -\dfrac{c+\delta+\sqrt{\Delta_1}}{2} < 0$ and $\lambda_2 = -\dfrac{c+\delta-\sqrt{\Delta_1}}{2} < 0$. Next we show that $\Delta_2 > \Delta_1$. Calculating the difference, we obtain

$$\begin{aligned}
\Delta_2 - \Delta_1 &= 4\epsilon_s c\delta - 4c\delta\left[1 - \frac{(1-\epsilon_r)\mathcal{R}_r}{(1-\mu)\mathcal{R}_s}\right] \\
&= 4c\delta(1-\epsilon_s)\left[\frac{(1-\epsilon_r)\mathcal{R}_r}{(1-\mu)(1-\epsilon_s)\mathcal{R}_s} - 1\right] \\
&= 4c\delta(1-\epsilon_s)\left[\frac{1}{(1-\mu)}\frac{\mathcal{R}_r'}{\mathcal{R}_s'} - 1\right],
\end{aligned}$$

where $\mathcal{R}_r' = (1-\epsilon_r)\mathcal{R}_r$ and $\mathcal{R}_s' = (1-\epsilon_s)\mathcal{R}_s$ 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, $\Delta_2 > \Delta_1$. Finally, we show that $(c+\delta)^2 > \Delta_2$. It suffices to show that $1 - \dfrac{(1-\epsilon_r)\mathcal{R}_r}{(1-\mu)\mathcal{R}_s} > 0$, i.e., $\dfrac{\mathcal{R}_r}{\mathcal{R}_s} < \dfrac{1-\mu}{1-\epsilon_r}$, 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. Therefore, $(c+\delta)^2 > \Delta_2 > \Delta_1 > (c-\delta)^2$. It follows that $\lambda_3 = -\dfrac{c+\delta+\sqrt{\Delta_2}}{2} < 0$ and $\lambda_4 = -\dfrac{c+\delta-\sqrt{\Delta_2}}{2} < 0$. Furthermore, from $\sqrt{\Delta_2} > \sqrt{\Delta_1}$, we have that $\lambda_3 < \lambda_1 < \lambda_2 < \lambda_4 < 0$.

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

(i) Notice that $C_1 = -\dfrac{c(1-2\epsilon_s)+\delta-\sqrt{\Delta_1}}{2\sqrt{\Delta_1}}V_s(0)$ and $\Delta_1 = (c+\delta)^2 - 4\epsilon_s c\delta$. $C_1 > 0$ is equivalent to $-c(1-2\epsilon_s) - \delta + \sqrt{\Delta_1} > 0$. Thus, for $C_1 > 0$ it suffices to prove that $\sqrt{\Delta_1} > c(1-2\epsilon_s) + \delta$. 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 $\Delta_1 > (c+\delta-2c\epsilon_s)^2$, which is equivalent to $\epsilon_s < 1$. Hence, $\Delta_1 > (c+\delta-2c\epsilon_s)^2$ and $C_1 > 0$.

(ii) Because $\sqrt{\Delta_1} > c - \delta$, we have

$$\begin{aligned}
c(1-2\epsilon_s) + \delta + \sqrt{\Delta_1} &> c(1-2\epsilon_s) + \delta + c - \delta \\
&= 2c(1-\epsilon_s) \\
&> 0.
\end{aligned}$$

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

(iii) For simplicity, we introduce a new parameter θ, defined as $\theta = 1 - \dfrac{1-\epsilon_r}{1-\mu}\dfrac{\mathcal{R}_r}{\mathcal{R}_s}$. Thus, $\theta < 1$.

Then C_3, C_4, and Δ_2 can be simplified to

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

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

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 $C_1e^{\lambda_1 t}$ and $C_2e^{\lambda_2 t}$ intersect, we obtain that

$$t_s = \frac{\ln\frac{C_1}{C_2}}{-\lambda_1+\lambda_2} = \frac{\ln\frac{C_1}{C_2}}{\sqrt{\Delta_1}}.$$

Similarly, we have $t_r = \frac{\ln\frac{C_3}{C_4}}{\sqrt{\Delta_2}}$. Calculating the difference between $\frac{C_1}{C_2}$ and $\frac{C_3}{C_4}$, we obtain

$$\frac{C_1}{C_2} - \frac{C_3}{C_4} = \frac{-c(1-2\epsilon_s)-\delta+\sqrt{\Delta_1}}{c(1-2\epsilon_s)+\delta+\sqrt{\Delta_1}} - \frac{-c(1-2\theta)-\delta+\sqrt{\Delta_2}}{c(1-2\theta)+\delta+\sqrt{\Delta_2}}, \tag{SI2-1}$$

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

Using the common denominator to combine two fractions in (SI2-1), we obtain the numerator

$$\Big[-c(1-2\epsilon_s)-\delta+\sqrt{\Delta_1}\Big]\Big[c(1-2\theta)+\delta+\sqrt{\Delta_2}\Big] - \Big[c(1-2\epsilon_s)+\delta+\sqrt{\Delta_1}\Big]\Big[-c(1-2\theta)-\delta+\sqrt{\Delta_2}\Big],$$

which can be simplified to

$$4c(\epsilon_s\sqrt{\Delta_2} - \theta\sqrt{\Delta_1}) - 2(c+\delta)(\sqrt{\Delta_2}-\sqrt{\Delta_1}). \tag{SI2-2}$$

Because drug resistant virus is more fit than wild-type virus during treatment, we have $(1-\epsilon_s)\mathcal{R}_s < (1-\epsilon_r)\mathcal{R}_r$. Thus, $\theta = 1 - \frac{1-\epsilon_r}{1-\mu}\frac{\mathcal{R}_r}{\mathcal{R}_s} < \epsilon_s$ since μ is very small. It follows from (SI2-2) that

$$\begin{aligned} 4c(\epsilon_s\sqrt{\Delta_2} - \theta\sqrt{\Delta_1}) - 2(c+\delta)(\sqrt{\Delta_2}-\sqrt{\Delta_1}) &> 4c(\epsilon_s\sqrt{\Delta_2} - \epsilon_s\sqrt{\Delta_1}) - 2(c+\delta)(\sqrt{\Delta_2}-\sqrt{\Delta_1}) \\ &= 2(\sqrt{\Delta_2}-\sqrt{\Delta_1})[2c\epsilon_s - (c+\delta)] \\ &> 0. \end{aligned}$$

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), $\theta = 1 - \dfrac{1-\epsilon_r}{1-\mu}\dfrac{\mathcal{R}_r}{\mathcal{R}_s}$ decreases and $\Delta_2 = (c+\delta)^2 - 4\theta c\delta$ increases. Rearranging $\dfrac{C_3}{C_4}$, 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

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

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, $\dfrac{C_3}{C_4}$ decreases and $t_r = \dfrac{\ln \frac{C_3}{C_4}}{\sqrt{\Delta_2}}$ decreases. This shows that t_r is an increasing function of ϵ_r.

Supporting Information 3: The Mutant Frequency in the Model with Backward Mutation

Before treatment, the model with backward mutation is

$$\begin{aligned}
\frac{d}{dt}T(t) &= s - dT - \beta_s V_s T - \beta_r V_r T,\\
\frac{d}{dt}I_s(t) &= \beta_s V_s T - \delta I_s,\\
\frac{d}{dt}I_r(t) &= \beta_r V_r T - \delta I_r,\\
\frac{d}{dt}V_s(t) &= (1-\mu)p_s I_s + \mu p_r I_r - cV_s,\\
\frac{d}{dt}V_r(t) &= \mu p_s I_s + (1-\mu)p_r I_r - cV_r.
\end{aligned} \tag{SI3-1}$$

There are only two possible steady states of the above model: 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

$$\mu p_r \beta_r \bar{T}\bar{V}_r = [c\delta - (1-\mu)p_s\beta_s\bar{T}]\bar{V}_s, \tag{SI3-2}$$

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

$$\mu p_s \beta_s \bar{T}\bar{V}_s = [c\delta - (1-\mu)p_r\beta_r\bar{T}]\bar{V}_r. \tag{SI3-3}$$

Thus, the two strains coexist only when

$$\frac{(1-\mu)p_s\beta_s\bar{T}}{c\delta} < 1 \tag{SI3-4}$$

and

$$\frac{(1-\mu)p_r\beta_r\bar{T}}{c\delta} < 1. \tag{SI3-5}$$

From (SI3-2) and (SI3-3), 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:

$$\bar{T}_{1,2} = \frac{(1-\mu)(p_s\beta_s + p_r\beta_r) \pm \sqrt{[(1-\mu)(p_s\beta_s + p_r\beta_r)]^2 - 4(1-2\mu)p_s\beta_s p_r\beta_r}}{2(1-2\mu)p_s\beta_s p_r\beta_r}c\delta. \tag{SI3-6}$$

Ignoring μ, we have two approximate solutions:

$$\bar{T}_1 \approx \frac{c\delta}{p_r\beta_r} \quad \text{(choosing "+" in (SI3-6))}$$

and

$$\bar{T}_2 \approx \frac{c\delta}{p_s\beta_s} \quad \text{(choosing "–" in (SI3-6)).}$$

Because of the conditions for the existence of the coexistence steady state, (SI3-4) and (SI3-5), only $\bar{T}_2$ is feasible. Thus, the mutant frequency before treatment can be calculated based on the equation (SI3-3):

$$\Phi = \frac{\bar{V}_r}{\bar{V}_s + \bar{V}_r} = \frac{1}{1 + \dfrac{c\delta - (1-\mu)p_r\beta_r\bar{T}_2}{\mu p_s\beta_s\bar{T}_2}},$$

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

$$\Phi = \frac{\mu}{\dfrac{(1-\mu)(1+r) + \sqrt{[(1-\mu)(1+r)]^2 - 4(1-2\mu)r}}{2} - r + \mu(1+r)}, \tag{SI3-7}$$

where $r = \dfrac{\mathcal{R}_r}{\mathcal{R}_s}$. It is clear that Φ depends only on μ and r.

It follows from (SI3-7) that Φ can be approximated by

$$\Phi = \frac{\mu}{1 - r + \mu(1+r)}, \tag{SI3-8}$$

which is less than $\dfrac{\mu}{1-r}$, the mutant frequency in the model without considering backward mutation. In fact, it can proved rigorously that $\Phi_w < \Phi_{wo}$, where Φ_w represents the mutant frequency with backward mutation (defined in (SI3-7)) and $\Phi_{wo} = \dfrac{\mu}{1-r}$ is the mutant frequency without backward mutation. For the proof, it suffices to show that

$$\frac{(1-\mu)(1+r) + \sqrt{[(1-\mu)(1+r)]^2 - 4(1-2\mu)r}}{2} + \mu(1+r) > 1.$$

The above 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. (SI3-8)), we observe that the difference between Φ_w and Φ_{wo} is miniscule. This shows that backward mutation does not play a significant role in the pre-treatment mutant frequency.

Supporting Information 4: The Pretreatment Mutant Frequency in the Model with Hepatocyte Proliferation

The model with hepatocyte proliferation is

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

From the V_s and V_r equations, we have

$$\bar{V}_s = \frac{(1-\mu)p_s\bar{I}_s}{c} \quad \text{and} \quad \bar{V}_r = \frac{\mu p_s\bar{I}_s + p_r\bar{I}_r}{c}.$$

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 $\bar{V}_s = \frac{(1-\mu)p_s\bar{I}_s}{c}$, 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 also depend on ρ_T, ρ_s, ρ_r, and other parameters.