Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(7): 12093–12117.
DOI: 10.3934/mbe.2023538
Received: 09 March 2023
Revised: 28 April 2023
Accepted: 08 May 2023
Published: 15 May 2023

Research article

The role of cell-to-cell transmission in HIV infection: insights from a mathematical modeling approach

Sophia Y. Rong[1], Ting Guo[2,4,*], J. Tyler Smith[4] and Xia Wang[3,4]

[1] Buchholz High School, Gainesville, FL 32606, USA

[2] Aliyun School of Big Data, Changzhou University, Changzhou 213164, China

[3] School of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China

[4] Department of Mathematics, University of Florida, Gainesville, FL 32611, USA

* **Correspondence:** Email: gt521521@cczu.edu.cn.

Abstract: HIV infection remains a serious global public health problem. Although current drug treatment is effective and can reduce plasma viral loads below the level of detection, it cannot eradicate the virus. The reasons for the low virus persistence despite long-term therapy have not been fully elucidated. In addition, multiple HIV infection, i.e., infection of a cell by multiple viruses, is common and can facilitate viral recombination and mutations, evading the immune system and conferring resistance to drug treatment. The mechanisms for multiple HIV infection formation and their respective contributions remain unclear. To answer these questions, we developed a mathematical modeling framework that encompasses cell-free viral infection and cell-to-cell spread. We fit sub-models that only have one transmission route and the full model containing both to the multi-infection data from HIV-infected patients, and show that the multi-infection data can only be reproduced if these two transmission routes are both considered. Computer simulations with the best-fitting parameter values indicate that cell-to-cell spread leads to the majority of multiple infection and also accounts for the majority of overall infection. Sensitivity analysis shows that cell-to-cell spread has reduced susceptibility to treatment and may explain low HIV persistence. Taken together, this work indicates that cell-to-cell spread plays a crucial role in the development of HIV multi-infection and low HIV persistence despite long-term therapy, and therefore has important implications for understanding HIV pathogenesis and developing more effective treatment strategies to control or even eliminate the disease.

Keywords: HIV; multiple infection; mathematical models; cell-to-cell spread

1. Introduction

As the causative agent of acquired immunodeficiency syndrome (AIDS), human immunodeficiency virus (HIV) infection remains a serious health problem in the world. At the end of 2021, there were approximately 38.4 million people living with HIV, about two-thirds of whom lived in Sub-Saharan Africa [1]. Since its discovery in 1981, the death toll of AIDS and the rapid spread of HIV have motivated a significant body of scientific and medical research. Over twenty antiretroviral drugs have been approved by the US Food and Drug Administration [2]. Currently, highly active antiretroviral therapy (HAART) containing three to five antiretroviral drugs has been successful in prolonging patients' survival and controlling viral replication. After 3–6 months of treatment, the viral load can be reduced below the limit of detection (i.e., 50 RNA copies per ml by standard assays), but supersensitive assays can reveal continual low-level viremia in patients under suppressive HAART [3, 4].

The mechanisms underlying low-level viral persistence are not yet fully understood. Earlier work [5, 6] suggested that viral persistence could be explained by the reactivation of cells from long-lived and stable latent reservoirs. Residual ongoing viral replication is also considered as a possible cause of viral persistence due to physiological sites or cellular compartments that have poor drug penetration [7, 8]. Zhang and Perelson [9] showed that the kinetics of leakage of virus from follicular dendritic cells (FDC) is consistent with some clinical data, suggesting that FDC might also contribute to low-level viral persistence.

HIV can spread through the population of target cells by two basic routes [10, 11]: (i) in cell-free virus infection, infected cells can produce free viruses, which infect susceptible target cells; (ii) in cell-to-cell spread, the virus can spread directly from an infected donor cell to a susceptible target cell, possibly through the formation of virological synapses [10–12]. There is experimental evidence that cell-to-cell spread is more effective than infection with free-floating viruses [13, 14], and it could play an important role in HIV dynamics. During cell-to-cell viral transmission, the probability that at least one of the virions escapes from antiretroviral inhibition increases as the number of transmitted virions increases [15]. Therefore, we will explore if cell-to-cell viral spread can explain the low HIV persistence during long-term therapy in this paper.

Cell-to-cell transmission can lead to the transfer of tens to hundreds of virus particles in a single synapse, a certain fraction of which successfully integrates into the target cell genome [12, 16], resulting in the multiple infection of cells. Multiply infected cells are frequently observed in experiments [17–19]. For example, Jung et al. [17] found that the HIV-infected splenocytes of two patients harbored between one and eight proviruses (a provirus is referred to as the HIV DNA integrated into the host cell's DNA), with an average of three to four proviruses per cell. During acute infection of macaques by simian immunodeficiency virus (SIV), infected CD4+ memory T cells carried an average of 1.5 viruses per cell [18]. Rebecca et al. [19] showed, using fluorescence in-situ hybridization, that an average of 3.4 proviral copies occurred per cell as a result of cell-to-cell viral transmission.

Mathematical models have been used to gain an understanding of HIV multiple infection [20–23]. Dixit and Perelson [20] first proposed probabilistic models of multiple infection and calculated the rate of infection when including only cell-free virus infection or both cell-free virus infection and cell-to-cell transmission. In a subsequent paper [21], they showed that the number of cells infected with i viruses correlates with the ith power of the number of singly infected cells under certain conditions. Ito et al. [22] developed an ordinary differential equation (ODE) model with the heterogeneity of the

target cell population, aimed at analyzing the HIV co-infection dynamics in the context of cell-free virus infection. They found that two different susceptible sub-populations were enough to reproduce the experimental data and around 98% of co-infected cells emerged from the more susceptible target cells. In a recent mathematical model of multiple infection [23], various numbers of virions were assumed to be spread with distinct probabilities.

Here we develop and examine a mathematical modeling framework incorporating multiple infected cell groups and two routes of viral spread. In order to investigate how multiple infection is formed, we compare the models with various scenarios of multiple infection to the HIV proviral copy data [17]. Simulating the model with the best-fit parameter estimates, we study the dynamics of infected cell populations, and evaluate the respective contributions of the two transmission routes to the singly and multiply infected cells. To determine if cell-to-cell viral spread can account for low HIV persistence in infected individuals receiving effective treatment, we also assess the sensitivity of the basic reproduction number and the level of HIV to the effectiveness of the drug.

2. Models and mathematical results

2.1. A model with two transmission routes

Based on basic models of within-host virus dynamics [24, 25], we formulate the following model incorporating cell-to-cell viral spread alongside cell-free viral infection.

$$\begin{cases} \dfrac{dT}{dt} = s - dT(t) - (1-\epsilon)\beta T(t)V(t) - \displaystyle\sum_{i=1}^{n} f_i(1-\epsilon_i)kT(t)\left(T^*(t) + \sum_{i=1}^{n} I_i(t)\right), \\ \dfrac{dT^*}{dt} = (1-\epsilon)\beta T(t)V(t) - \delta T^*(t), \\ \dfrac{dI_i}{dt} = f_i(1-\epsilon_i)kT(t)\left(T^*(t) + \displaystyle\sum_{j=1}^{n} I_j(t)\right) - \delta_i I_i(t), \quad i = 1, 2, \ldots, n, \\ \dfrac{dV}{dt} = N\left(\delta T^*(t) + \displaystyle\sum_{i=1}^{n} \delta_i I_i(t)\right) - cV(t). \end{cases} \tag{2.1}$$

In the model, T and V denote the concentrations of uninfected CD4+ T cells and free virus, respectively. T^* denotes the concentration of CD4+ T cells infected by free virions, and I_i ($i = 1, 2, \ldots, n$) represents cells infected through cell-to-cell spread with i viruses transmitted. Uninfected cells are generated at rate s, and die at rate d. β and k represent the infection rates of uninfected cells via direct cell-free viral infection and cell-to-cell spread, respectively. The parameter f_i ($i = 1, 2, \ldots, n$) denotes the chance that cell-to-cell spread transmits i viral genomes. Because the sum of all the possibilities is equal to 1, we have that $\sum_{i=1}^{n} f_i = 1$. Infected cells die at the rate δ or δ_i, corresponding to their respective class. It is assumed that virions are generated at a rate $N\delta$ from the cells infected by the cell-free virus, and at rate $N\delta_i$ from cells infected by cell-to-cell spread, where N is the viral burst size, that is the total amount of virions created by one infected cell during its lifetime (i.e., $1/\delta$ for T^* and $1/\delta_i$ for I_i). Free virions are assumed to be removed from circulation at the rate c. The parameter ϵ represents the drug's effectiveness in inhibiting cell-free virus infection, and ϵ_i ($i = 1, 2, \ldots, n$) is the effectiveness of antiretroviral treatment in inhibiting the infection by i virions via cell-to-cell transmission. Sigal et al. [15] suggested that when more virions are spread in each cell-to-cell transmission, the therapy is

less likely to inhibit the infection induced by all the transmitted virions. Therefore, it is assumed that $\epsilon_1 > \epsilon_2 > \cdots > \epsilon_n$. The schematic diagram of the model (2.1) is shown in Figure 1, and all parameters and their values are summarized in Table 1.

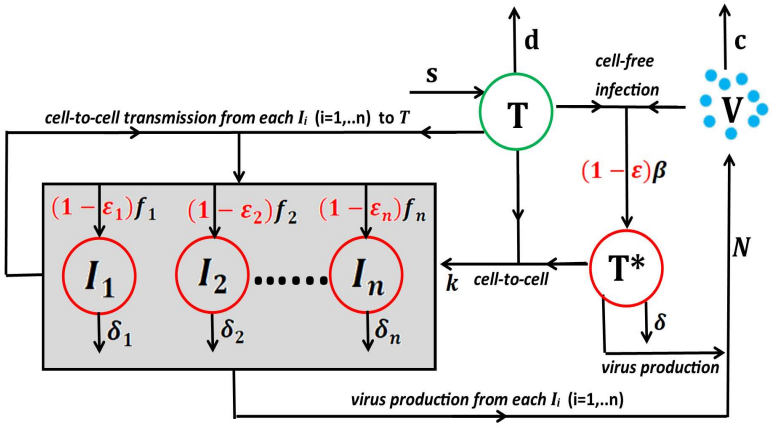

Figure 1. Schematic illustration of Eq (2.1) incorporating two routes of viral spread. Variables T, T^*, I_i and V denote uninfected CD4+ T cells, cells infected by the free virus, cells infected via cell-to-cell spread that transmits i virions, and free virus, respectively. The parameter definitions and values are listed in Table 1.

Table 1. Parameters and values used in the mathematical model (2.1).

Parameter	Description	Value	Source
s	Generation rate of uninfected CD4+ T cells	10^4 ml^{-1}day^{-1}	[23]
d	Death rate of uninfected CD4+ T cells	0.01 day^{-1}	[23,26]
β	Infection rate of cells by cell-free virus	Data fitting	See text
k	Cell-to-cell viral transmission rate	Data fitting	See text
f_i	Probability of cell-to-cell transmission that sends the cell to class I_i	Data fitting	See text
ϵ	Drug's effect of inhibiting cell-free virus infection	See text	See text
ϵ_i	Drug's effect of inhibiting cell-to-cell viral spread	See text	See text
δ	Death rate of infected cells T^*	1 day^{-1}	[23,27,28]
δ_i	Death rate of infected cells I_i	See text	See text
N	Viral burst size	Data fitting	See text
c	Viral clearance rate	23 day^{-1}	[29]

2.2. *Model analysis*

We will use model (2.1) to analyze the HIV multiple infection distribution data in two patients who did not take any antiretroviral therapy [17]. Thus, we first let all the drug efficacies be 0, i.e., $\epsilon = \epsilon_i = 0$, for $i = 1, 2, \ldots, n$. The model with drug treatment will be discussed later. The model without antiretroviral therapy is given by

$$\begin{cases} \dfrac{dT}{dt} = s - dT(t) - \beta T(t)V(t) - kT(t)\left(T^*(t) + \sum\limits_{i=1}^{n} I_i(t)\right), \\ \dfrac{dT^*}{dt} = \beta T(t)V(t) - \delta T^*(t), \\ \dfrac{dI_i}{dt} = f_i kT(t)T^*(t) + f_i kT(t)\sum\limits_{j=1}^{n} I_j(t) - \delta_i I_i(t), \quad i = 1, 2, \ldots, n, \\ \dfrac{dV}{dt} = N\left(\delta T^*(t) + \sum\limits_{i=1}^{n} \delta_i I_i(t)\right) - cV(t). \end{cases} \tag{2.2}$$

Any steady state of model (2.2) must satisfy the following equalities

$$\begin{cases} s - dT - \beta TV - kT\left(T^* + \sum\limits_{i=1}^{n} I_i\right) = 0, \\ \beta TV - \delta T^* = 0, \\ f_i kTT^* + f_i kT\sum\limits_{j=1}^{n} I_j - \delta_i I_i = 0, \quad i = 1, 2, \cdots, n, \\ N\left(\delta T^* + \sum\limits_{i=1}^{n} \delta_i I_i\right) - cV = 0. \end{cases} \tag{2.3}$$

It is clear that model (2.2) always admits an infection-free equilibrium $E_0 = (T_0, 0, 0, \cdots, 0)$ with $T_0 = s/d$, where there is no infection.

Using the next-generation method [30], we derive the basic reproduction number, given by (see Appendix A).

$$R_0 = \frac{1}{2}\left(\frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i} + \sqrt{\left(\frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i}\right)^2 + 4\frac{N\beta}{c}\sum_{i=1}^{n} k\left(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}\right)}\right)T_0. \tag{2.4}$$

The basic reproduction number is a key measure in estimating the ability of a disease to spread in epidemiological research. It is defined as the average number of secondary transmissions from one infected person in a wholly susceptible environment. When it is greater than 1, the epidemic is growing; when it is less than 1, the disease is predicted to die out. In the context of the within-host viral dynamic model (2.2), R_0 represents the average number of newly infected cells (or virions) generated by one infected cell (or one virion) introduced into an environment where all cells are susceptible. The complexity of the expression, in particular the square root, makes it hard to explain the biological meaning. However, with straightforward algebraic computation, we find that the following R_0' is an equivalent measure but it is easy to explain the biological meaning.

$$R_0' = \frac{N\beta}{c}T_0 + kT_0\sum_{i=1}^{n}\frac{f_i}{\delta_i} + \frac{N\beta T_0}{c}\sum_{i=1}^{n} k\left(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}\right)T_0 \tag{2.5}$$

It can be easily shown that $R_0 > 1$, $R_0 = 1$, and $R_0 < 1$ are equivalent to $R_0' > 1$, $R_0' = 1$, and $R_0' < 1$, respectively. R_0' consists of three parts, each representing the contribution of infection by free virus,

by cell-to-cell transmission from the infected cell populations I_i, and by cell-to-cell spread from the infected cell population T^*.

In addition to the infection-free equilibrium E_0, model (2.2) has an infected equilibrium $E_1 = (\hat{T}, \hat{T}^*, \hat{I}_1, \hat{I}_2, \ldots, \hat{I}_n, \hat{V})$, where

$$\begin{aligned}
\hat{T} &= \frac{c}{N\beta}\frac{\left(\frac{N\beta s}{c\delta} - \left(\frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i}\right)\hat{T}^*\right)}{\frac{d}{\delta} + k\sum_{i=1}^{n}\left(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}\right)\hat{T}^*} \\
\hat{I}_j &= \frac{\hat{T}\hat{T}^*}{1 - \hat{T}k\sum_{i=1}^{n}\frac{f_i}{\delta_i}}k\frac{f_j}{\delta_j}, \quad j = 1, 2, \ldots, n \\
\hat{V} &= \frac{N\delta}{c}\left(1 + \frac{\hat{T}}{\delta}\frac{k}{1 - \hat{T}k\sum_{i=1}^{n}\frac{f_i}{\delta_i}}\right)\hat{T}^*
\end{aligned} \tag{2.6}$$

and $\hat{T}^*$ is the unique positive solution of the following equation

$$\frac{k(Ns\beta - q_2\hat{T}^*)^2}{cN\beta\delta(d + \delta q_1\hat{T}^*)^2(1 - k\sum_{i=1}^{n}\frac{f_i}{\delta_i}\frac{Ns\beta - q_2\hat{T}^*}{N\beta(d+\delta q_1\hat{T}^*)})} = 1 - \frac{Ns\beta - q_2\hat{T}^*}{c(d + \delta q_1\hat{T}^*)},$$

with

$$q_1 = k\sum_{i=1}^{n}\left(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}\right), \qquad q_2 = \frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i}.$$

To study the local stability of the infection-free equilibrium E_0 of model (2.2), we linearize model (2.2) and obtain the following Jacobian matrix evaluated at E_0

$$J(E_0) = \begin{pmatrix}
-d & -kT_0 & -kT_0 & -kT_0 & \ldots & -kT_0 & -kT_0 & -\beta T_0 \\
0 & -\delta & 0 & 0 & \ldots & 0 & 0 & \beta T_0 \\
0 & kf_1T_0 & kf_1T_0 - \delta_1 & kf_1T_0 & \ldots & kf_1T_0 & kf_1T_0 & 0 \\
0 & kf_2T_0 & kf_2T_0 & kf_2T_0 - \delta_2 & \ldots & kf_2T_0 & kf_2T_0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & kf_{n-1}T_0 & kf_{n-1}T_0 & kf_{n-1}T_0 & \ldots & kf_{n-1}T_0 - \delta_{n-1} & kf_{n-1}T_0 & 0 \\
0 & kf_nT_0 & kf_nT_0 & kf_nT_0 & \ldots & kf_nT_0 & kf_nT_0 - \delta_n & 0 \\
0 & N\delta & N\delta_1 & N\delta_2 & \ldots & N\delta_{n-1} & N\delta_n & -c
\end{pmatrix}. \tag{2.7}$$

Similarly, we can obtain the corresponding Jacobian matrix of model (2.2) at the infected equilib-

rium E_1, given by

$$J(E_1) = \begin{pmatrix} -\frac{(\beta\hat{V}+d)\left(\hat{T}k\sum_{i=1}^{n}\frac{f_i}{\delta_i}-1\right)-\hat{T}^*k}{\hat{T}k\sum_{i=1}^{n}\frac{f_i}{\delta_i}-1} & -k\hat{T} & -k\hat{T} & -k\hat{T} & \dots & -k\hat{T} & -\beta\hat{T} \\ \beta\hat{V} & -\delta & 0 & 0 & \dots & 0 & \beta\hat{T} \\ -\frac{f_1k\hat{T}^*}{k\hat{T}\sum_{i=1}^{n}\frac{f_i}{\delta_i}-1} & kf_1\hat{T} & kf_1\hat{T}-\delta_1 & kf_1\hat{T} & \dots & kf_1\hat{T} & 0 \\ -\frac{f_2k\hat{T}^*}{k\hat{T}\sum_{i=1}^{n}\frac{f_i}{\delta_i}-1} & kf_2\hat{T} & kf_2\hat{T} & kf_2\hat{T}-\delta_2 & \dots & kf_2\hat{T} & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{f_nk\hat{T}^*}{k\hat{T}\sum_{i=1}^{n}\frac{f_i}{\delta_i}-1} & kf_n\hat{T} & kf_n\hat{T} & kf_n\hat{T} & \dots & kf_n\hat{T}-\delta_n & 0 \\ 0 & N\delta & N\delta_1 & N\delta_2 & \dots & N\delta_n & -c \end{pmatrix}. \tag{2.8}$$

With distinct death rates of infected cells, the analysis of the model for local stability is challenging. However, with the assumption of equal death rates (i.e., $\delta = \delta_1 = \cdots = \delta_n$), the infection-free equilibrium E_0 and new infection matrix $\mathcal{F}_{n+2}$ remain unchanged, and the basic reproduction number becomes $R_0 = \frac{N\beta}{c}T_0 + \frac{k}{\delta}T_0$ (Appendix A). In this case, when $R_0 > 1$, there exists an infected equilibrium $E_1 = (\hat{T}, \hat{T}^*, \hat{I}_1, \hat{I}_2, \dots, \hat{I}_n, \hat{V})$, where

$$\begin{aligned} \hat{T} &= \frac{c\delta}{N\delta\beta + kc}, & \hat{T}^* &= \frac{c\delta\beta Nd(R_0-1)}{(N\delta\beta+kc)^2}, \\ \hat{V} &= \frac{\delta Nd(R_0-1)}{N\delta\beta+kc}, & \hat{I}_i &= \frac{c^2dkf_i(R_0-1)}{(N\delta\beta+kc)^2}, \quad i = 1, 2, \dots, n. \end{aligned} \tag{2.9}$$

It is clear that the infected equilibrium exists if and only if $R_0 > 1$. For the local stability of the infection-free equilibrium E_0 and the infected equilibrium E_1, we have the following results, with the proof given in Appendixes B and C, respectively.

Theorem 2.1. The infection-free equilibrium E_0 is locally asymptotically stable if $R_0 < 1$ and unstable if $R_0 > 1$.

Theorem 2.2. The infected equilibrium E_1 is locally asymptotically stable whenever it exists, i.e., when $R_0 > 1$.

3. Data analysis and computer simulations

3.1. Multiple infection data analysis

Model (2.2) will be used to analyze the multiple HIV infection data in Jung et al. [17]. The data showed a bimodal frequency distribution for the number of HIV-1 proviral copies per infected cell in two patients who were not on antiretroviral therapy. The proviral copy number was between one and eight per cell, with a mean of about 3.2. Thus, we have $n = 8$ in our model. We calculate the fraction of cells containing i proviruses in the overall infection when the system is at equilibrium (because the two patients were in the chronic stage of infection), and the proportion is denoted by ψ_i ($i = 1, 2, \cdots, n$).

We also reorganize the data by dividing the number of cells containing a given number of proviruses by the total number of cells collected from each patient [17]. In this way, we obtain the frequency of multiple infection and compare it with our model prediction.

Because of the limited frequency data in these two patients, we keep the assumption that all the infected cells have the same death rate. We also set some parameter values based on existing research [23,26–29,31]. In the absence of infection, CD4+ T cell levels are approximately $10^6\ ml^{-1}$ [31]. The death rate of uninfected cells d is $0.01\ day^{-1}$ [23,26]. Thus, the generation rate of uninfected cells s is $10^6 \times 0.01 = 10^4\ ml^{-1}day^{-1}$. The death rate of infected cells δ is fixed to $1\ day^{-1}$, which is the same as in [23,27]. The clearance rate of virus c is chosen to be $23\ day^{-1}$ [28,29]. We fit the calculated frequency ψ_i (which will be derived later for each sub-model) to the proviral distribution data of each patient using the nonlinear least square method and provide estimates for the other parameters.

Multi-infection may be formed by cell-to-cell transmission originating from the directly virus-infected cells T^*, cell-to-cell transmission from the cells I_i, or a combination of both. In order to discriminate the potential mechanism of multiple infection formation and determine the respective contribution from each transmission route, we consider the following two sub-models and the full model.

Sub-model A: Only virus-directly-infected cells transmit viruses

If only the virus-directly-infected cells T^* transit viruses, then the sub-model is given by

$$\begin{cases} \dfrac{dT}{dt} = s - dT(t) - \beta T(t)V(t) - kT(t)T^*(t), \\ \dfrac{dT^*}{dt} = \beta T(t)V(t) - \delta T^*(t), \\ \dfrac{dI_i}{dt} = f_i kT(t)T^*(t) - \delta I_i(t), \quad i = 1, 2, \ldots, n, \\ \dfrac{dV}{dt} = N\delta\left(T^*(t) + \displaystyle\sum_{i=1}^{n} I_i(t)\right) - cV(t). \end{cases} \tag{3.1}$$

This simplified sub-model A has an infected equilibrium $E_A = (T_A, T_A^*, I_{1A}, I_{2A}, \ldots, I_{nA}, V_A)$, where

$$\begin{aligned} T_A &= -\frac{\delta}{k} + \sqrt{(\frac{\delta}{k})^2 + \frac{4c\delta}{N\beta}}, \qquad V_A = \frac{Ns}{c} - \frac{Nd}{c}T_A, \\ T_A^* &= \frac{\beta}{\delta}T_A V_A, \qquad I_{iA} = \frac{kf_i}{\delta}T_A T_A^*, \quad i = 1, 2, \ldots, n. \end{aligned} \tag{3.2}$$

According to this equilibrium, we compute the fraction of infected cells containing i proviruses, given by

$$\psi_1 = \frac{I_{1A} + T_A^*}{\sum_{i=1}^{n} I_{iA} + T_A^*}, \qquad \psi_2 = \frac{I_{2A}}{\sum_{i=1}^{n} I_{iA} + T_A^*}, \quad \cdots, \quad \psi_8 = \frac{I_{8A}}{\sum_{i=1}^{n} I_{iA} + T_A^*}$$

The chance of spreading i virions during cell-to-cell transmission (f_i, $i = 1, 2, \cdots, 8$) remains unknown. Inspired by the work of Dixit and Perelson [20], we first assume that f_i takes the Poisson

distribution $f_i = \frac{e^{-\theta}\theta^i}{i!}$, where θ is the mean value of viral genomes spread in each cell-to-cell transmission. Thus, $\sum_{i=0}^{\infty} \frac{e^{-\theta}\theta^i}{i!} = 1$, which leads to $\sum_{i=1}^{\infty} \frac{e^{-\theta}\theta^i}{i!} = 1 - e^{-\theta}$, i.e., $\sum_{i=1}^{\infty} \frac{e^{-\theta}}{1-e^{-\theta}} \frac{\theta^i}{i!} = 1$. We modify the Poisson distribution to $f_i = \frac{e^{-\theta}}{1-e^{-\theta}} \frac{\theta^i}{i!}$ to rule out the event that cell-to-cell transmission does not transmit any virus. The data (blue circles) and the best fit are displayed in Figure 2 upper panels. For both patients, the model fitting is similar in the distribution of proviral copy numbers. Only singly infected cells exist (Figure 2 upper panels). This is not consistent with the data that one cell has a mean of 3.2 proviruses [17]. Thus, model (3.1) with a modified Poisson distribution of f_i fails to reproduce the multiple infection data [17].

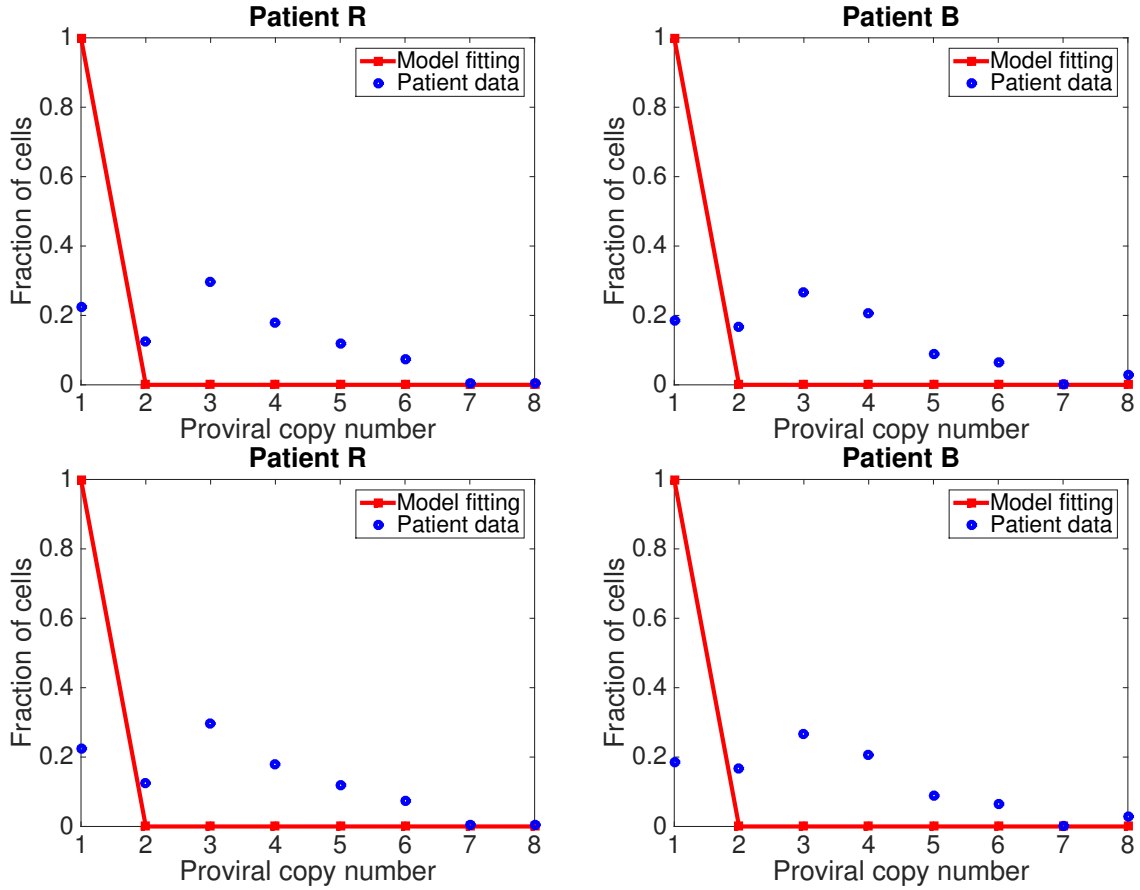

Figure 2. Model fitting to the multiple infection data (blue circle). Here we assume there is only cell-to-cell spread originating from virus-directly-infected cells T^*. Upper panels: the chance f_i of spreading i virions follows a Poisson probability distribution. Low panels: the probability follows the binomial distribution. The model including only viral spread from T^* cannot explain the multiple infection data.

We also use other distributions of the probability f_i in the data fitting. In Figure 2 lower panels, we use a binomial distribution, which was also used in some other models [16, 32, 33]. We assume the chance of spreading i genomes is

$$f_i = \binom{n}{i} \gamma^i (1-\gamma)^{n-i},$$

where γ is the probability of having a successful infection from one virus. Using the binomial distribution, we have the same fitting (Figure 2 lower panels). Therefore, model (3.1) cannot explain the patient multiple infection data using either a Poisson distribution or binomial distribution. We also try some other distributions such as Gaussian and log-normal distributions and get similar fitting results.

Sub-model B: Only cell-to-cell spread transmits viruses

In this sub-model, we assume there is no direct cell-free virus infection and there is only cell-to-cell transmission. The sub-model is given by

$$\begin{cases} \dfrac{dT}{dt} = s - dT(t) - kT(t)\displaystyle\sum_{i=1}^{n} I_i(t), \\ \dfrac{dI_i}{dt} = f_i kT(t)\displaystyle\sum_{j=1}^{n} I_j(t) - \delta I_i(t), \quad i = 1, 2, \ldots, n, \\ \dfrac{dV}{dt} = N\delta\displaystyle\sum_{i=1}^{n} I_i(t) - cV(t). \end{cases} \tag{3.3}$$

The steady state of infected cells with i proviruses is $I_{iB} = f_i(\frac{s}{\delta} - \frac{d}{k})$ and its frequency in the total infected cell population is given by

$$\psi_i = \frac{I_{iB}}{\sum_{j=1}^{n} I_{jB}}, \quad i = 1, 2, \ldots, n.$$

We compare the frequency with the same data set [17], as shown in Figure 3 upper panels. It was found that model (3.3) with the Poisson distribution of f_i provides a reasonable data fitting, except for missing the first peak (i.e., infected cells with one proviral copy). Using a binomial distribution, the model also provides a reasonable fit to the data (see Figure 3 lower panels). Therefore, we conclude that model (3.3) assuming only cell-to-cell spread without direct cell-free virus infection provides a good fit to the patient data. Next, we investigate which distribution is better for reproducing the multiple infection data. For the comparison of the fitting using the two distributions, we calculate SSR (see below), the sum of squared residuals, which measures the error between model prediction and experimental data. We also compute AIC, the Akaike information criterion, which also evaluates the relative quality of different mathematical models in fitting a given set of data.

$$\text{SSR} = \sum_{i=1}^{n}(\psi_i - \hat{\psi}_i)^2, \qquad \text{AIC} = n\ln(\text{SSR}/n) + 2\sigma, \tag{3.4}$$

where n is the number of data points from each patient, σ is the number of parameters used in data fitting, ψ_i is the predicted frequency of infected cells containing i proviral copies calculated from our model and $\hat{\psi}_i$ is the corresponding patient data. Based on the best-fitting results, we present the SSR and AIC values in Table 2. It turns out that model (3.3) with the Poisson distribution of f_i ($model_{im}$ of Table 2) has smaller SSR and AIC than the binomial distribution (i.e., $model_{ib}$ of Table 2). This indicates that the model with Poisson distribution is better at fitting the data.

Table 2. Model comparison*.

	SSR				AIC			
Patient	$Model_{im}$	$Model_{ib}$	$Model_{bm}$	$Model_{bb}$	$Model_{im}$	$Model_{ib}$	$Model_{bm}$	$Model_{bb}$
R	0.021	0.028	0.018	0.007	-43.54	-41.24	-40.77	-48.33
B	0.009	0.014	0.008	0.003	-50.32	-46.79	-47.26	-55.11

* Models are compared by SSR and AIC, which are calculated in Eq (3.4).
$Model_{im}$ is Eq (3.3) with Poisson distribution;
$Model_{ib}$ is Eq (3.3) with binomial distribution;
$Model_{bm}$ is Eq (3.5) with Poisson distribution;
$Model_{bb}$ is Eq (3.5) with binomial distribution.

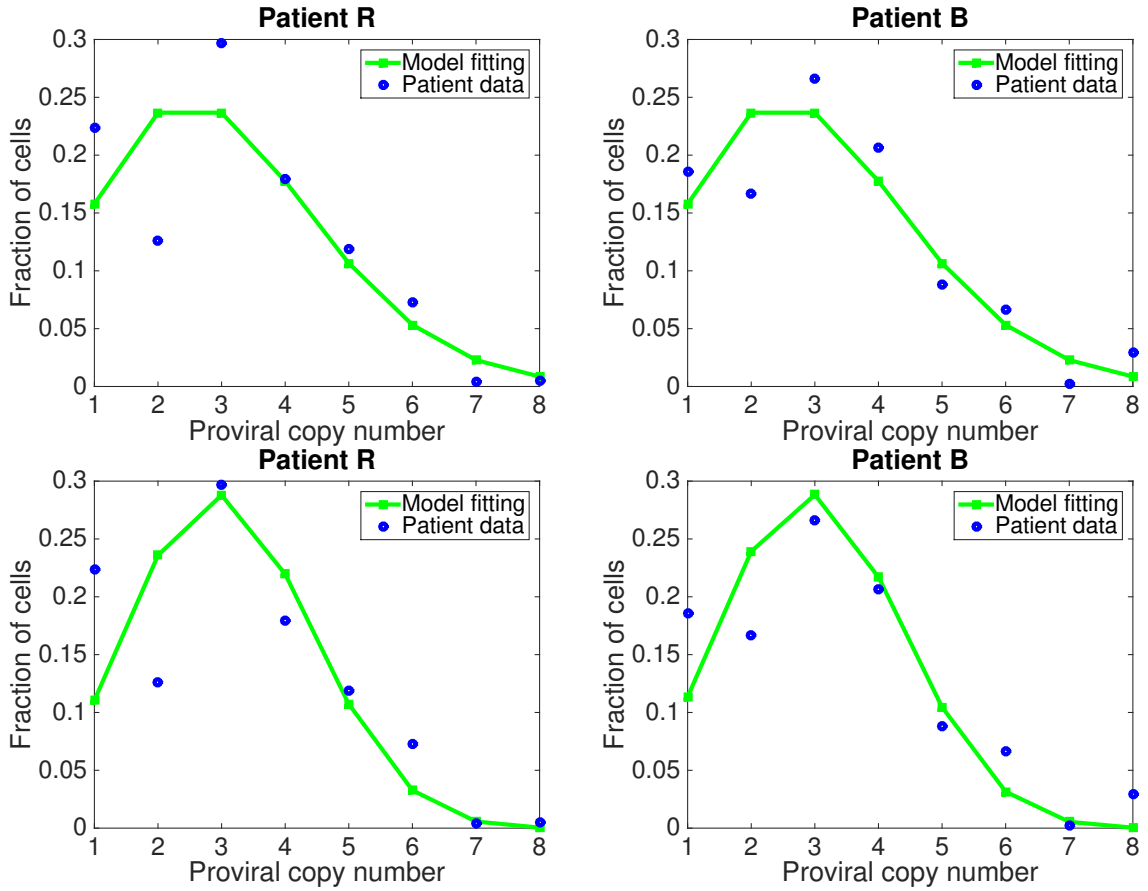

Figure 3. Model fitting to experimental data. Here we assume that there is no direct cell-free virus infection (i.e., model (3.3)). The model prediction agrees with the patient data. The upper panels use Poisson distribution and the lower panels use binomial distribution.

Although the model prediction with only cell-to-cell spread (Eq (3.3)) agrees with the patient data, the fitting did not capture the bimodal proviral peaks at one and three proviral copies. We tried some other distributions, e.g., Gaussian distribution and log-normal distribution, but found that they cannot explain the two peaks either. The best fitting always shows a peak for infected cells containing three proviruses (similar to the fitting shown in Figure 3). Because the model assuming only virus-directly-infected cells can transmit viruses (i.e., Eq (3.1)) always exhibits a single peak

for singly infected cells in the best fitting (Figure 2), we speculate that both transmission routes are required to capture the two-peak distribution of multi-infection. In the following, we use the full model to fit the same data set and further evaluate the relative contributions from each transmission mode.

Full model C: Both transmission routes can transmit viruses

To study if both the two transmission routes are required to capture the two-peak distribution of proviruses, we use the following full model

$$\begin{cases} \dfrac{dT}{dt} = s - dT(t) - \beta T(t)V(t) - kT(t)\left(T^*(t) + \sum\limits_{i=1}^{n} I_i(t)\right), \\ \dfrac{dT^*}{dt} = \beta T(t)V(t) - \delta T^*(t), \\ \dfrac{dI_i}{dt} = f_i k T(t)T^*(t) + f_i k T(t)\sum\limits_{j=1}^{n} I_j(t) - \delta I_i(t), \quad i = 1, 2, \ldots, n, \\ \dfrac{dV}{dt} = N\delta\left(T^*(t) + \sum\limits_{i=1}^{n} I_i(t)\right) - cV(t). \end{cases} \tag{3.5}$$

For the full model, the infected steady state E_1 is given in Eq (2.9). The frequency of infected cells containing i proviruses is calculated from this steady state as follow:

$$\psi_1 = \frac{I_1 + T^*}{\sum_{i=1}^{n} I_i + T^*}, \ \ \psi_2 = \frac{I_2}{\sum_{i=1}^{n} I_i + T^*}, \ \ \cdots, \psi_8 = \frac{I_8}{\sum_{i=1}^{n} I_i + T^*}.$$

Table 3. Best-fitting parameter estimates.

Patient	β (ml/day)	k (ml/day)	γ (no unit)	N ($virus/cell$)
R	1.65×10^{-8}	5.63×10^{-6}	0.42	1447
B	1.2×10^{-8}	4.73×10^{-6}	0.41	1170
Mean	1.425×10^{-8}	5.18×10^{-6}	0.415	1308.5
SD	2.25×10^{-9}	4.5×10^{-7}	0.005	138.5

Figure 4 shows the model fitting of the calculated frequency to the multiple infection distribution data. The upper panels use the modified Poisson distribution and the lower panels use the binomial distribution for the probability f_i. We find that both can capture the bimodal distribution at one and three proviral copies. Therefore, model (3.5) improves the fitting to the proviral distribution data. For a further comparison of the data fitting using the full model, SSR and AIC values were calculated for each distribution (see Table 2). The full model with the binomial distribution, denoted by $model_{bb}$, has smaller SSR and AIC values than $model_{bm}$, the model (3.5) with the Poisson distribution.

In conclusion, we find that the two routes of viral spread are both required to capture the two-peak data of multiple infection shown in experiment [17]. Model comparison from Table 2 suggests that the full model (3.5) with a binomial distribution for cell-to-cell transmission yields the smallest SSR and AIC values. The best-fitting parameter values using the full model and binomial distribution are provided in Table 3.

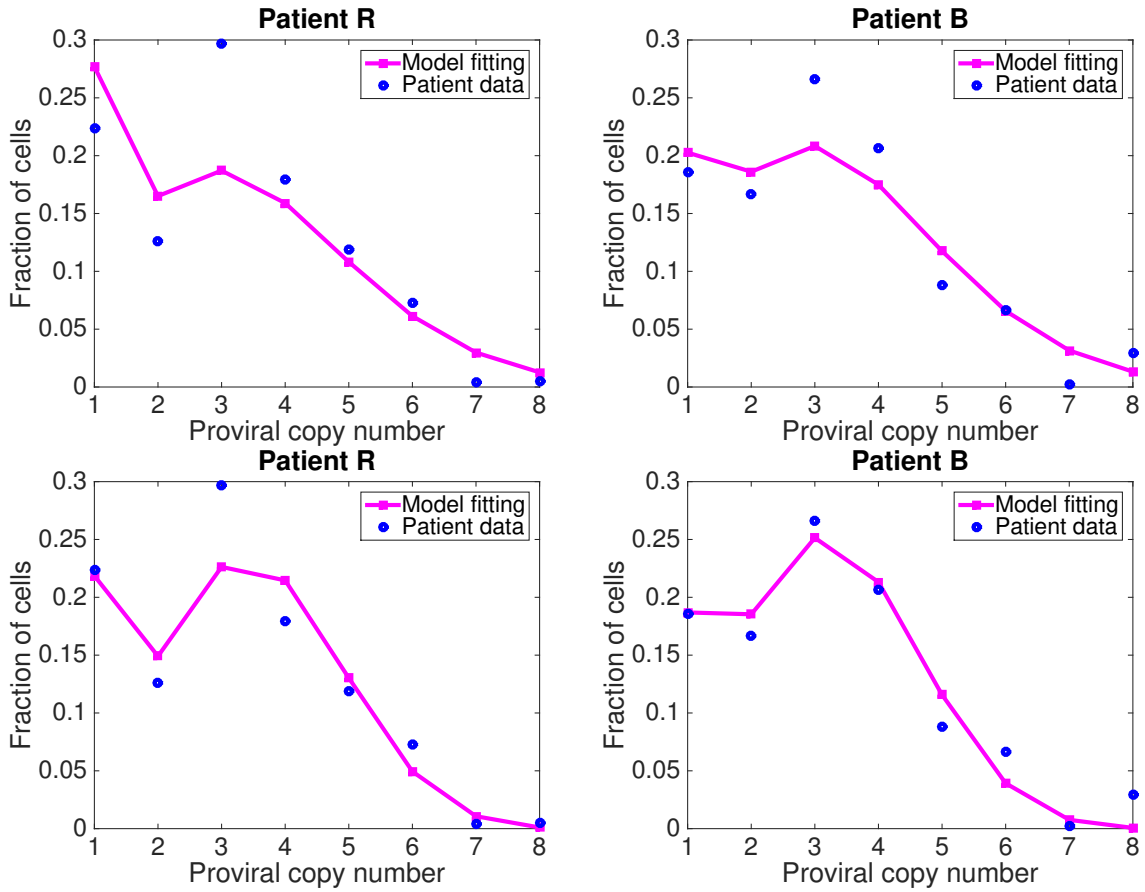

Figure 4. Data fitting using the full model incorporating cell-to-cell viral spread from both T^* and all of I_i. The fitting can reproduce the bimodal multiple-cell distribution at one and three proviral copies.

3.2. Respective contributions to multiple infection

Data fitting to the multiple infection data in the last section suggests that the two transmission routes are both required to explain the two-peak proviral distribution. In this section, we evaluate the relative contribution of these two transmission routes to the multiple infection and total infection. Using the best-fitting parameter values of the full model, we simulate the dynamics of all of the infected cell populations. Figure 5(a) shows that all the infected cells containing a different number of proviruses have similar dynamics despite different magnitudes. For each class of I_i, the cell population increases rapidly, then decreases, oscillates, and eventually converges to a steady state. Using a different model, Dixit and Perelson [21] found that the abundance of the infected cell subpopulation at equilibrium scales with the number of proviruses present in cells, i.e., singly infected cells are most abundant, followed by cells containing two, three, or four viruses, etc. However, in our prediction, the cells containing three proviruses are most abundant, followed by cells infected with four, one, two, five, six proviruses, etc (see the zoom-in figure in Figure 5(a)). Cell-to-cell transmission that is included in our model may explain the difference.

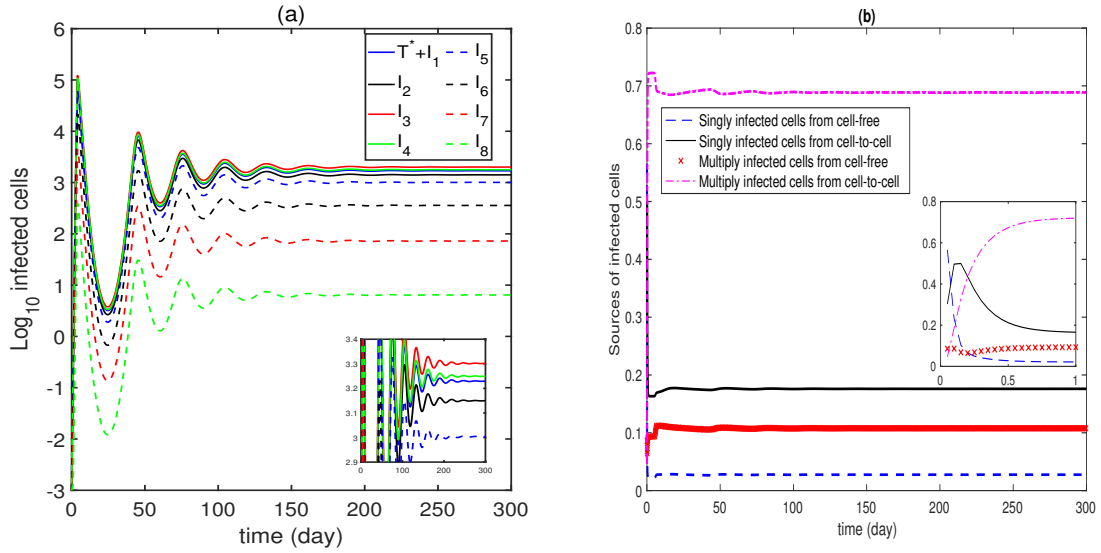

Figure 5. Simulation of the full model with best-fitting parameter values from Table 3. (a) Predicted time evolution of cells containing different amounts of proviral copies. (b) The respective contribution to the overall infection. They include single and multiple infection, as well as cell-free virus infection and cell-to-cell transmission.

To elaborate on the contribution to the total infection from single infection and multiple infection, we separate infected cells into two categories, single infection $T^*(t) + I_1(t)$ and multiple infection $\sum_{i=2}^{8} I_i(t)$. The infection can be established either from cell-free virus infection or cell-to-cell spread. The relative contributions from these two transmission routes are given by

$$\phi_1 = \frac{\beta T(t)V(t)}{\beta T(t)V(t) + kT(t)\left(T^*(t) + \sum_{i=1}^{8} I_i(t)\right)} \quad \text{and} \quad \phi_2 = \frac{kT(t)\left(T^*(t) + \sum_{i=1}^{8} I_i(t)\right)}{\beta T(t)V(t) + kT(t)\left(T^*(t) + \sum_{i=1}^{8} I_i(t)\right)}.$$

Therefore, the contribution to the total infection from single infection and from cell-free virus infection, from single infection and from cell-to-cell spread, from multiple infection and from cell-free virus infection, and from multiple infection and from cell-to-cell spread is respectively given by

$$\phi_1 \cdot \frac{T^*(t) + I_1(t)}{T^*(t) + \sum_{i=1}^{8} I_i(t)}, \quad \phi_2 \cdot \frac{T^*(t) + I_1(t)}{T^*(t) + \sum_{i=1}^{8} I_i(t)}, \quad \phi_1 \cdot \frac{\sum_{i=2}^{8} I_i(t)}{T^*(t) + \sum_{i=1}^{8} I_i(t)}, \quad \phi_2 \cdot \frac{\sum_{i=2}^{8} I_i(t)}{T^*(t) + \sum_{i=1}^{8} I_i(t)}.$$

Using the same best-fitting parameter values as in Figure 5(a), we plot the above relative contributions in Figure 5(b). We observe that at the beginning of infection, cells infected with a single virus from cell-free virus infection are the most common (zoomed-in figure in Figure 5(b)). As the infection continues, the contributions to the total infection from multi-infection and from cell-to-cell spread

increase rapidly. All curves have some minor fluctuation before converging to the respective steady states. Multiple infection from cell-to-cell transmission accounts for the majority of the total infected cells (about 70%). The contribution from multiple infection initially via cell-free virus infection accounts for approximately 10% of the total infection. The remaining infection is singly infected cells, in which cell-to-cell spread and direct viral infection account for 18 and 2%, respectively. This shows that cell-to-cell viral spread plays a crucial role in developing multi-infection and also accounts for most cell infection. The percentages may be slightly different due to the potential variation in parameter estimates. However, this should not affect the major conclusion in view of their relative magnitude.

3.3. Influence of cell-to-cell spread on HIV persistence during therapy

To determine if cell-to-cell spread can explain the low HIV persistence despite long-term effective antiretroviral treatment, we study the viral change when drug efficacy increases. We start with the sensitivity of the basic reproduction number during treatment, denoted by R_0^T, with respect to the drug effectiveness against direct viral infection ϵ and the drug effectiveness against cell-to-cell spread ϵ_i ($i = 1, 2, \ldots, 8$).

In the presence of therapy, the basic reproduction number given in (2.4) becomes

$$R_0^T = \frac{1}{2}\Bigg(\sqrt{\left(\frac{N(1-\epsilon)\beta}{c} + k\sum_{i=1}^{n}(1-\epsilon_i)\frac{f_i}{\delta_i}\right)^2 + 4\frac{N(1-\epsilon)\beta}{c}\sum_{i=1}^{n}k(1-\epsilon_i)\left(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}\right)} + \frac{N(1-\epsilon)\beta}{c} + k\sum_{i=1}^{n}(1-\epsilon_i)\frac{f_i}{\delta_i}\Bigg)T_0.$$

For each drug efficacy, we calcualte the normalized forward sensitivity index of R_0^T as

$$\begin{aligned} \frac{\partial R_0^T}{\partial \epsilon}\Big/\frac{R_0^T}{\epsilon} &= -\frac{N\beta\epsilon}{c}\frac{1}{p_1 + \sqrt{p_2}}\left(1 + \frac{1}{\sqrt{p_2}}(p_1 + 2\sum_{i=1}^{n}k(1-\epsilon_i)(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}))\right), \\ \frac{\partial R_0^T}{\partial \epsilon_i}\Big/\frac{R_0^T}{\epsilon_i} &= -\frac{k\epsilon_i}{p_1 + \sqrt{p_2}}\left(\frac{f_i}{\delta_i} + \frac{1}{\sqrt{p_2}}(\frac{p_1 f_i}{\delta_i} + \frac{2(1-\epsilon)N\beta}{c}(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}))\right), \quad i = 1, 2, \ldots, 8, \end{aligned} \tag{3.6}$$

where

$$p_1 = \frac{N\beta(1-\epsilon)}{c} + \sum_{i=1}^{n}\frac{k(1-\epsilon_i)f_i}{\delta_i},$$

$$p_2 = p_1^2 + 4\frac{N\beta(1-\epsilon)}{c}\sum_{i=1}^{n}k(1-\epsilon_i)(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}).$$

Using the best-fitting parameter estimates in Table 3, we plot the sensitivity index in Figure 6(a). All the indices are less than zero, indicating a negative correlation between the basic reproduction number and drug efficacy. The basic reproduction number is relatively sensitive to the drug effectiveness against cell-free virus infection (i.e., ϵ). However, the basic reproduction number is not sensitive to some drug

effectiveness against cell-to-cell spread, such as ϵ_6, ϵ_7, and ϵ_8. This shows that cell-to-cell viral spread may explain the low viral load persistence during therapy, particularly when the drug effectiveness against cell-free virus infection is fixed.

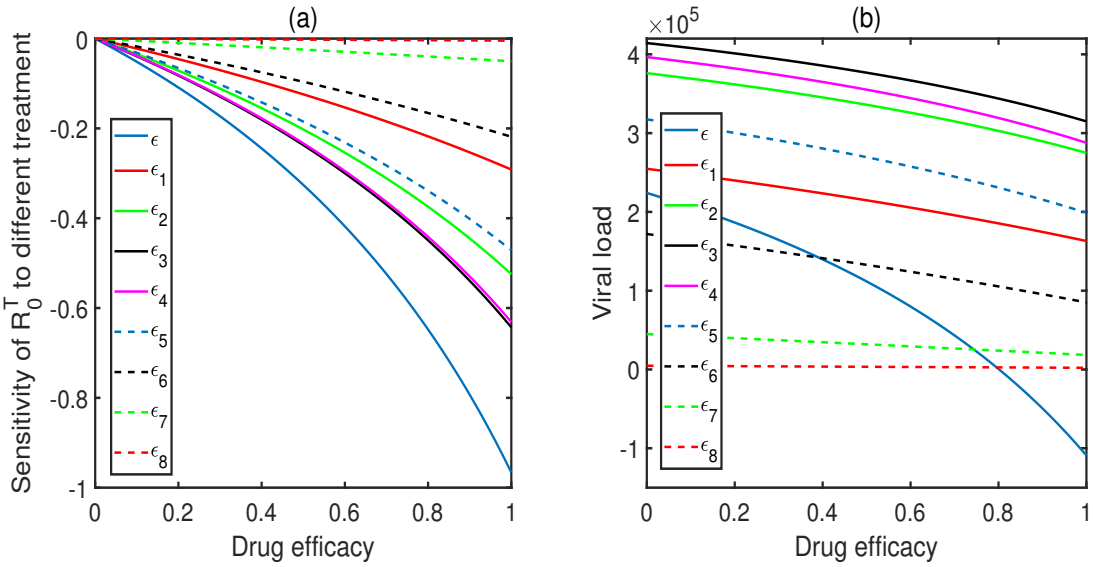

Figure 6. Sensitivity tests of (a) the basic reproduction number R_0^T and (b) the final viral load to drug's effectiveness ϵ and ϵ_i ($i = 1, 2, \ldots, 8$). The values of the parameters except that used for the sensitivity test are from Figure 5.

We further study how the steady-state viral load changes as the drug efficacy increases. Like the sensitivity of the basic reproduction number, the viral load is relatively responsive to the alteration in the drug effectiveness against cell-free virus infection. However, it is not sensitive to the drug's effectiveness against cell-to-cell spread. (see Figure 6(b)). Taken together, these analyses show that the viral load is not sensitive to some drug efficacies against cell-to-cell transmission. This suggests that cell-to-cell viral spread can explain low HIV persistence despite long-term combination treatment.

4. Conclusions and discussion

Increasing evidence from in vivo and in vitro studies suggests the high prevalence of HIV multi-infection [17–19,34,35]. However, the biological factors responsible for HIV multiple infection remain unclear, especially the two-peak distribution data in [17]. In this study, we construct a mathematical modeling framework including both infection from cell-free viral infection and cell-to-cell viral spread to explain HIV multiple infection. By fitting mathematical models to the distribution data of multiple infection, we discover that there are only single infected cells if the model only includes cell-to-cell spread from T^* (Figure 2). Including cell-to-cell transmission alone without cell-free viral infection can generate the peak for cells containing three proviruses (see Figure 3). The results in Figure 4 show that both the infected cells T^* and I_i are required in our model to explain the two-peak proviral distribution. This suggests that the two modes of transmission both contribute to the formation of HIV multi-infection.

We did not include sequential cell-free virus infection in the model because the cell can down-regulate its surface CD4 molecules following the initial infection, rendering further infections difficult [36, 37]. However, the CD4 down-modulation takes about one day, during which infected cells still remain susceptible to further infections [38]. In this case, a certain amount of co-infection takes place and may lead to multiple infection. The exact shape of the provirus distribution and which cell population is most abundant may depend on the choice of the rate of receptor down-modulation [39]. When the rate is relatively slow, two peaks occur in the distribution: one for cells infected with a single virus, and one for cells infected with four viruses (see Figure 7c in [39]). The distribution can be altered for different rates of receptor down-modulation, as demonstrated in [39]. The sequential virus infection with a decreasing infection rate due to CD4 down-regulation was included in a model and it was shown that the sequential virus infection alone cannot explain the multiple infection [40].

From some experiments and modeling work [15, 33], cell-to-cell viral transmission, e.g., via the formation of virological synapses between cells, renders the infection less susceptible to antiretroviral drugs. In our work, we assume that the effect of treatment in inhibiting cell-to-cell viral spread decreases as the number of transmitted virions during cell-to-cell spread increases. Sensitivity analyses of the basic reproduction number and the viral load during therapy show that the model prediction is less sensitive to the drug effectiveness of inhibiting cell-to-cell spread, although it is still sensitive to the drug effectiveness of blocking cell-free viral infection (Figure 6). Therefore, cell-to-cell spread may play a critical role in explaining the low HIV persistence during suppressive therapy.

Some studies showed that the number of single infected cells is higher than the total number of multiply infected cells [21, 41]. However, in our work and also in the experiment [17], the population of multiply infected cells is more frequent than the population of single infected cells. This difference can be attributed to the spread of the virus from one cell to another. In in vitro experiments and their modeling studies [21, 41], the number of virions is much greater than the number of cells by several orders of magnitude. Therefore, cell-free viral infection dominates and cell-to-cell viral spread may be neglected. However, in in vivo experiments such as Jung et al. [17] and this study, viral transmission from cell to cell is highly possible. Model predictions based on the data on multiple cell infections in [17] show that more than 80% of the total infection is due to cell-to-cell virus transmission. This viral mode of transmission also leads to about 70% of multiple infections (Figure 5(b)), further suggesting that cell-to-cell viral spread plays a critical role in the development of HIV multi-infection in vivo.

Another reason we introduce cell-to-cell viral transmission is that infected CD4 T cells are isolated from the spleens of two individuals [17], where the cells are relatively densely packed and have a high probability of synapsing. In contrast, the cells in the blood mix better and may inhibit the formation of synaptic connections. This can lead to the frequent single infection of target cells [42]. The concentration of target cells is higher in the lymphatic tissues than in the blood, so cells may be infected several times either through sequential viral infection [39] or cell-to-cell spread discussed here. In summary, the routes of spread of the virus and thus the types of infected cells (single or multiply infected cells) may also vary depending on the tissue site.

HIV recombination is a crucial consequence of multiple infection. Recombination between genetically different viruses in the same infected cell has been observed in the spleens of HIV-infected patients [17]. Although the effect of recombination on HIV evolution is complex, the integration of different alleles into a single genome can confer drug resistance to antiretroviral drugs and facilitate evasion of the host immune responses [43–47]. The differential equation system employed here to

examine the dynamics of multiple HIV infections may not be suitable for investigating HIV recombination, as this typically involves the use of bioinformatics tools and phylogenetic methods to analyze genetic sequence data and reconstruct the evolutionary relationships between different HIV lineages.

Elucidating the mechanisms underlying multiple infections and HIV persistence has important implications for understanding HIV pathogenesis and developing more effective treatment strategies to control infection. Our modeling study, combined with experimental data [17], shows that cell-to-cell virus transmission can play an important role in the development of multiple infection and HIV persistence despite long-term effective combination antiretroviral therapy. Therefore, all treatment strategies that can effectively block cell-to-cell transmission, e.g., by blocking the establishment of viral synapses that might be required for cell-to-cell transmission, would be crucial for the control or even elimination of the disease. This work also provides a novel mathematical modeling framework that can examine some other issues in HIV infection, such as HIV latent infection, which is considered another important obstacle to viral elimination in current antiretroviral treatment. HIV can establish latent infection very rapidly following initial infection, and the resulting reservoir of virus remains relatively stable during therapy but can be activated to release new viral particles. If this latent reservoir accounts for the majority of the persistent virus, it could explain the lack of viral evolution observed in patients during suppressive therapy [48, 49].

Acknowledgments

T. Guo was supported by the Changzhou Scientific and Technological Program grant (CJ20220134), the Natural Science Foundation of Jiangsu Higher Education (22KJB110007), and the National Natural Science Foundation of China (12201077, 12071217). X. Wang was supported by the National Natural Science Foundation of China (No. 12171413), the Natural Science Foundation of Henan Province (222300420016), and the Program for Innovative Research Team (in Science and Technology) in Universities of Henan Province (21IRTSTHN014).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. WHO, HIV/AIDS: key facts, 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/hiv-aids.
2. FDA Approval of HIV medicines, AIDSinfo, 2018. Available from: https://aidsinfo.nih.gov/understanding-hiv-aids/infographics/25/fda-approval-of-hiv-medicines.
3. G. Dornadula, H. Zhang, B. VanUitert, J. Stern, L. Livornese, M. J. Ingerman, et al., Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy, *JAMA*, **282** (1999), 1627–1632. https://doi.org/10.1001/jama.282.17.1627
4. F. Maldarelli, Targeting viral reservoirs: ability of antiretroviral therapy to stop viral replication, *Curr. Opin. HIV AIDS*, **6** (2011), 49–56. https://doi.org/10.1097/COH.0b013e32834134ea

5. S. Palmer, F. Maldarelli, A. Wiegand, B. Bernstein, G. J. Hanna, S. C. Brun, et al., Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy, *Proc. Natl. Acad. Sci. USA*, **105** (2008), 3879–3884. https://doi.org/10.1073/pnas.0800050105
6. L. Shan, R. F. Siliciano, From reactivation of latent HIV-1 to elimination of the latent reservoir: the presence of multiple barriers to viral eradication, *BioEssays*, **35** (2013), 544–552. https://doi.org/10.1002/bies.201200170
7. N. Ching, O. O. Yang, J. G. Deville, K. Nielsen-Saines, B. J. Ank, M. S. Sim, et al., Pediatric HIV-1-specific cytotoxic T-lymphocyte responses suggesting ongoing viral replication despite combination antiretroviral therapy, *Pediatr. Res.*, **61** (2007), 692–697. https://doi.org/10.1203/pdr.0b013e31805365ef
8. T. Guo, Z. Qiu, L. Rong, Modeling the role of macrophages in HIV persistence during antiretroviral therapy, *J. Math. Biol.*, **81** (2020), 369–402. https://doi.org/10.1007/s00285-020-01513-x
9. J. Zhang, A. S. Perelson, Contribution of follicular dendritic cells to persistent HIV viremia, *J. Virol.*, **87** (2013), 7893–7901. https://doi.org/10.1128/JVI.00556-13
10. J. Feldmann, O. Schwartz, HIV-1 virological synapse: live imaging of transmission, *Viruses*, **2** (2010), 1666–1680. https://doi.org/10.3390/v2081666
11. Q. J. Sattentau, Cell-to-cell spread of retroviruses, *Viruses*, **2** (2010), 1306–1321. https://doi.org/10.3390/v2061306
12. W. Hübner, G. McNerney, P. Chen, B. M. Dale, R. E. Gordon, F. Y. S. Chuang, et al., Quantitative 3D video microscopy of HIV transfer across T cell virological synapses, *Science*, **323** (2009), 1743–1747. https://doi.org/10.1126/science.1167525
13. L. M. Agosto, P. Uchil, W. Mothes, HIV cell-to-cell transmission: effects on pathogenesis and antiretroviral therapy, *Trends Microbiol.*, **23** (2015), 289–295. https://doi.org/10.1016/j.tim.2015.02.003
14. P. Zhong, L. M. Agosto, J. B. Munro, W. Mothes, Cell-to-cell transmission of viruses, *Curr. Opin. Virol.*, **3** (2013), 44–50. https://doi.org/10.1016/j.coviro.2012.11.004
15. A. Sigal, J. T. Kim, A. B. Balazs, E. Dekel, A. Mayo, R. Milo, et al., Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy, *Nature*, **477** (2011), 95–98. https://doi.org/10.1038/nature10347
16. N. L. Komarova, D. Wodar, Virus dynamics in the presence of synaptic transmission, *Math. Biosci.*, **242** (2013), 161–171. https://doi.org/10.1016/j.mbs.2013.01.003
17. A. Jung, R. Maier, J. P. Vartanian, G. Bocharov, V. Jung, U. Fischer, et al., Recombination: multiply infected spleen cells in HIV patients, *Nature*, **418** (2002), 144. https://doi.org/10.1038/418144a
18. J. J. Mattapallil, D. C. Douek, B. Hill, Y. Nishimura, M. Martin, M. Roederer, Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection, *Nature*, **434** (2005), 1093–1097. https://doi.org/10.1038/nature03501
19. A. Rebecca, M. Nicola, M. Ivonne, E. Jones, Q. J. Sattentau, Multiple proviral integration events after virological synapse-mediated HIV-1 spread, *Virology*, **443** (2013), 143–149. https://doi.org/10.1016/j.virol.2013.05.005

20. N. M. Dixit, A. S. Perelson, Multiplicity of human immunodeficiency virus infections in lymphoid tissue, *J. Virol.*, **78** (2004), 8942–8945. https://doi.org/10.1128/JVI.78.16.8942-8945.2004

21. N. M. Dixit, A. S. Perelson, HIV dynamics with multiple infections of target cells, *Proc. Natl. Acad. Sci. USA*, **102** (2005), 8198–8203. https://doi.org/10.1073/pnas.0407498102

22. Y. Ito, A. Tauzin, A. Remion, K. Ejima, F. Mammano, S. Iwami, Dynamics of HIV-1 coinfection in different susceptible target cell populations during cell-free infection, *J. Theor. Biol.*, **455** (2018), 39–46. https://doi.org/10.1016/j.jtbi.2018.06.025

23. X. Wang, L. Rong, HIV low viral load persistence under treatment: insights from a model of cell-to-cell viral transmission, *Appl. Math. Lett.*, **94** (2019), 44–51. https://doi.org/10.1016/j.aml.2019.02.019

24. A. S. Perelson, P. W. Nelson, Mathematical analysis of HIV-1 dynamics in vivo, *SIAM Rev.*, **41** (1999), 3–44. https://doi.org/10.1137/S0036144598335107

25. M. A. Nowak, R. M. May, *Virus Dynamics: Mathematical Principles of Immunology and Virology*, Oxfors: Oxford University Press, 2000.

26. H. Mohri, S. Bonhoeffer, S. Monard, A. S. Perelson, D. D. Ho, Rapid turnover of T lymphocytes in SIV-infected rhesus macaques, *Science*, **279** (1998), 1223–1227. https://doi.org/10.1126/science.279.5354.1223

27. L. Rong, A. Perelson, Modeling HIV persistence, the latent reservoir, and viral blips, *J. Theor. Biol.*, **260** (2009), 308–331. https://doi.org/10.1016/j.jtbi.2009.06.011

28. T. Guo, Z. Qiu, The effects of CTL immune response on HIV infection model with potent therapy, latently infected cells and cell-to-cell viral transmission, *Math. Biosci. Eng.*, **16** (2019), 6822–6841. https://doi.org/10.3934/mbe.2019341

29. X. Wang, S. Tang, X. Song, L. Rong, Mathematical analysis of an HIV latent infection model including both virus-to-cell infection and cell-to-cell transmission, *J. Biol. Dyn.*, **11** (2017), 455–483. https://doi.org/10.1080/17513758.2016.1242784

30. P. V. den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48. https://doi.org/10.1016/S0025-5564(02)00108-6

31. M. Bofill, G. Janossy, C. A. Lee, D. Macdonald-Burns, A. N. Phillips, C. Sabin, et al., Laboratory control values for CD4 and CD8 T lymphocytes. Implications for HIV-1 diagnosis, *Clin. Exp. Immunol.*, **88** (1992), 243–252. https://doi.org/10.1111/j.1365-2249.1992.tb03068.x

32. N. L. Komarova, D. N. Levy, D. Wodarz, Effect of synaptic transmission on viral fitness in HIV infection, *PLoS One*, **7** (2012), e48361. https://doi.org/10.1371/journal.pone.0048361

33. N. Komarova, D. Levy, D. Wodarz, Synaptic transmission and the susceptibility of HIV infection to anti-viral drugs, *Sci. Rep.*, **3** (2013), 2103. https://doi.org/10.1038/srep02103

34. M. Boullé, T. Müller, S. Dähling, Y. Ganga, L. Jackson, D. Mahamed, et al., HIV cell-to-cell spread results in earlier onset of viral gene expression by multiple infections per cell, *PLoS Pathog.*, **12** (2016), e1005964. https://doi.org/10.1371/journal.ppat.1005964

35. A. R. Templeton, M. G. Kramer, J. Jarvis, J. Kowalski, S. Gange, M. F. Schneider, et al., Multiple-infection and recombination in HIV-1 within a longitudinal cohort of women, *Retrovirology*, **6** (2009), 54. https://doi.org/10.1186/1742-4690-6-54

36. J. Lama, The physiological relevance of CD4 receptor down-modulation during HIV infection, *Curr. HIV Res.*, **1** (2003), 167–184. https://doi.org/10.2174/1570162033485276

37. K. Levesque, A. Finzi, J. Binette, E. A. Cohen, Role of CD4 receptor down-regulation during HIV-1 infection, *Curr. HIV Res.*, **2** (2004), 51–59. https://doi.org/10.2174/1570162043485086

38. M. Nethe, B. Berkhout, A. C. van der Kuyl, Retroviral superinfection resistance, *Retrovirology*, **2** (2005), 52. https://doi.org/10.1186/1742-4690-2-52

39. D. Wodarz, D. N. Levy, Effect of different modes of viral spread on the dynamics of multiply infected cells in human immunodeficiency virus infection, *J. R. Soc. Interface*, **8** (2011), 289–300. https://doi.org/10.1098/rsif.2010.0266

40. T. Guo, Z. Qiu, K. Kitagawa, S. Iwami, L. Rong, Modeling HIV multiple infection, *J. Theor. Biol.*, **509** (2021), 110502. https://doi.org/10.1016/j.jtbi.2020.110502

41. D. N. Levy, G. M. Aldrovandi, O. Kutsch, G. M. Shaw, Dynamics of HIV-1 recombination in its natural target cells, *Proc. Natl. Acad. Sci. USA*, **101** (2004), 4204–4209. https://doi.org/10.1073/pnas.0306764101

42. L. Josefsson, M. King, B. Makitalo, J. Brännström, W. Shao, F. Maldarelli, et al., Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule, *Proc. Natl. Acad. Sci. USA*, **108** (2011), 11199–11204. https://doi.org/10.1073/pnas.1107729108

43. H. Song, E. E. Giorgi, V. V. Ganusov, F. Cai, G. Athreya, H. Yoon, et al., Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection, *Nat. Commun.*, **9** (2018), 1928. https://doi.org/10.1038/s41467-018-04217-5

44. A. Carvajal-Rodríguez, K. A. Crandall, D. Posada, Recombination favors the evolution of drug resistance in HIV-1 during antiretroviral therapy, *Infect. Genet. Evol.*, **7** (2007), 476–483. https://doi.org/10.1016/j.meegid.2007.02.001

45. J. Rawson, O. A. Nikolaitchik, B. F. Keele, V. K. Pathak, Recombination is required for efficient HIV-1 replication and the maintenance of viral genome integrity, *Nucleic Acids Res.*, **46** (2018), 10535–10545. https://doi.org/10.1093/nar/gky910

46. J. Kreger, N. L. Komarova, D. Wodarz, Effect of synaptic cell-to-cell transmission and recombination on the evolution of double mutants in HIV, *J. R. Soc. Interface*, **17** (2020), 20190832. https://doi.org/10.1098/rsif.2019.0832

47. M. Arenas, N. M. Araujo, C. Branco, N. Castelhano, E. Castro-Nallar, M. Pérez-Losada, Mutation and recombination in pathogen evolution: relevance, methods and controversies, *Infect. Genet. Evol.*, **63** (2018), 295–306. https://doi.org/10.1016/j.meegid.2017.09.029

48. M. Kearney, J. Spindler, W. Shao, S. Yu, E. M. Anderson, A. O'Shea, et al., Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy, *PLoS Pathog.*, **10** (2014), e1004010. https://doi.org/10.1371/journal.ppat.1004010

49. G. Bozzi, F. R. Simonetti, S. A. Watters, E. M. Anderson, M. Gouzoulis, M. F. Kearney, et al., No evidence of ongoing HIV replication or compartmentalization in tissues during combination antiretroviral therapy: implications for HIV eradication, *Sci. Adv.*, **5** (2019), eaav2045. https://doi.org/10.1126/sciadv.aav2045

Appendix

A. Derivation of the basic reproduction number

Using the next-generation method [30], we define the new infection matrix $\mathcal{F}_{n+2}$ and the transfer matrix $\mathcal{V}_{n+2}$, given by the following expressions. The infection matrix in an epidemiological model is a matrix that describes the rates at which individuals in different compartments of the model (e.g., susceptible, infected, recovered) become infected when they come into contact with individuals in other compartments. The transfer matrix, on the other hand, is a matrix that describes the rates at which individuals move between compartments in the model.

$$\mathcal{F}_{n+2} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & \beta T_0 \\ f_1 k T_0 & f_1 k T_0 & f_1 k T_0 & \cdots & f_1 k T_0 & f_1 k T_0 & 0 \\ f_2 k T_0 & f_2 k T_0 & f_2 k T_0 & \cdots & f_2 k T_0 & f_2 k T_0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ f_{n-1} k T_0 & f_{n-1} k T_0 & f_{n-1} k T_0 & \cdots & f_{n-1} k T_0 & f_{n-1} k T_0 & 0 \\ f_n k T_0 & f_n k T_0 & f_n k T_0 & \cdots & f_n k T_0 & f_n k T_0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{V}_{n+2} = \begin{pmatrix} \delta & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \delta_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \delta_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \delta_{n-1} & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \delta_n & 0 \\ -N\delta & -N\delta_1 & -N\delta_2 & \cdots & -N\delta_{n-1} & -N\delta_n & c \end{pmatrix}.$$

It is clear that $\mathcal{F}_{n+2}$ is non-negative and $\mathcal{V}_{n+2}$ is non-singular. Thus, the basic reproduction number can be calculated by the spectral radius of the next generation matrix $\mathcal{M} = \mathcal{F}_{n+2}\mathcal{V}_{n+2}^{-1}$, given by

$$\mathcal{M} = \begin{pmatrix} \frac{\beta N}{c}T_0 & \frac{\beta N}{c}T_0 & \frac{\beta N}{c}T_0 & \cdots & \frac{\beta N}{c}T_0 & \frac{\beta N}{c}T_0 & \frac{\beta}{c}T_0 \\ \frac{f_1 k}{\delta}T_0 & \frac{f_1 k}{\delta_1}T_0 & \frac{f_1 k}{\delta_2}T_0 & \cdots & \frac{f_1 k}{\delta_{n-1}}T_0 & \frac{f_1 k}{\delta_n}T_0 & 0 \\ \frac{f_2 k}{\delta}T_0 & \frac{f_2 k}{\delta_1}T_0 & \frac{f_2 k}{\delta_2}T_0 & \cdots & \frac{f_2 k}{\delta_{n-1}}T_0 & \frac{f_2 k}{\delta_n}T_0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \frac{f_{n-1} k}{\delta}T_0 & \frac{f_{n-1} k}{\delta_1}T_0 & \frac{f_{n-1} k}{\delta_2}T_0 & \cdots & \frac{f_{n-1} k}{\delta_{n-1}}T_0 & \frac{f_{n-1} k}{\delta_n}T_0 & 0 \\ \frac{f_n k}{\delta}T_0 & \frac{f_n k}{\delta_1}T_0 & \frac{f_n k}{\delta_2}T_0 & \cdots & \frac{f_n k}{\delta_{n-1}}T_0 & \frac{f_n k}{\delta_n}T_0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}. \tag{A.1}$$

Therefore, the basic reproduction number is

$$\begin{aligned} R_0 &= \rho(\mathcal{M}) \\ &= \frac{1}{2}\left(\frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i} + \sqrt{\left(\frac{N\beta}{c} - k\sum_{i=1}^{n}\frac{f_i}{\delta_i}\right)^2 + 4\frac{N\beta k}{c\delta}} \right) T_0 \\ &= \frac{1}{2}\left(\frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i} + \sqrt{\left(\frac{N\beta}{c} + k\sum_{i=1}^{n}\frac{f_i}{\delta_i}\right)^2 + 4\frac{N\beta}{c}\sum_{i=1}^{n} k\left(\frac{f_i}{\delta} - \frac{f_i}{\delta_i}\right)} \right) T_0. \end{aligned} \tag{A.2}$$

B. Proof of Theorem 2.1

From the Jacobinan matrix (2.7), we obtain the characteristic equation for the infection-free steady state E_0 as follows:

$$|\lambda I - J(E_0)| = \begin{vmatrix} \lambda + d & kT_0 & kT_0 & kT_0 & \cdots & kT_0 & \beta T_0 \\ 0 & \lambda + \delta & 0 & 0 & \cdots & 0 & -\beta T_0 \\ 0 & -f_1 kT_0 & \lambda - f_1 kT_0 + \delta & -f_1 kT_0 & \cdots & -f_1 kT_0 & 0 \\ 0 & -f_2 kT_0 & -f_2 kT_0 & \lambda - f_2 kT_0 + \delta & \cdots & -f_2 kT_0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & -f_n kT_0 & -f_n kT_0 & -f_n kT_0 & \cdots & \lambda - f_n kT_0 + \delta & 0 \\ 0 & -N\delta & -N\delta & -N\delta & \cdots & -N\delta & \lambda + c \end{vmatrix} = 0, \tag{B.1}$$

where λ is the eigenvalue and I is the identity matrix with the same dimension as $J(E_0)$. Expanding (B.1), we obtain

$$(\lambda + d)(\lambda + \delta)^n\Big(\lambda^2 + (\frac{N\beta\delta s}{cd} + (1 - R_0)\delta + c)\lambda + c\delta(1 - R_0)\Big) = 0. \tag{B.2}$$

Equation (B.2) always has a negative root $\lambda = -d$ and the negative root $\lambda = -\delta$ with multiplicity n. Thus, the local stability of E_0 is determined by the roots of following equation

$$\lambda^2 + \Big(\frac{N\beta\delta s}{cd} + (1 - R_0)\delta + c\Big)\lambda + c\delta(1 - R_0) = 0. \tag{B.3}$$

Let

$$F(\lambda) = \lambda^2 + \Big(\frac{N\beta\delta s}{cd} + (1 - R_0)\delta + c\Big)\lambda + c\delta(1 - R_0).$$

When $R_0 > 1$, $F(\lambda) = 0$ has at least one positive real root. When $R_0 < 1$, all roots of $F(\lambda) = 0$ have negative real parts. Thus, the infection-free equilibrium E_0 is locally asymptotically stable when $R_0 < 1$ and unstable when $R_0 > 1$. This completes the proof.

C. Proof of Theorem 2.2

From the Jacobian matrix (2.8), we obtain the following characteristic equation at the infected steady state E_1

$$|\lambda I - J(E_1)| = \begin{vmatrix} \lambda + d + \beta\hat{V} + k(\hat{T}^* + \sum_{i=1}^{n}\hat{I}_i) & k\hat{T} & k\hat{T} & \cdots & k\hat{T} & \beta\hat{T} \\ -\beta\hat{V} & \lambda + \delta & 0 & \cdots & 0 & -\beta\hat{T} \\ -f_1k(\hat{T}^* + \sum_{i=1}^{n}\hat{I}_i) & -f_1k\hat{T} & \lambda - f_1k\hat{T} + \delta & \cdots & -f_1k\hat{T} & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -f_nk(\hat{T}^* + \sum_{i=1}^{n}\hat{I}_i) & -f_nk\hat{T} & -f_nk\hat{T} & \cdots & \lambda - f_nk\hat{T} + \delta & 0 \\ 0 & -N\delta & -N\delta & \cdots & -N\delta & \lambda + c \end{vmatrix} = 0, \tag{C.1}$$

where λ is the eigenvalue, and $\hat{T}, \hat{T}^*, \hat{I}_i$ and $\hat{V}$ are given in Eq (2.9). Expanding the above determinant and regrouping terms, we obtain the following characteristic equation

$$(\lambda + \delta)^n(\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3) = 0, \tag{C.2}$$

where

$$\begin{aligned} a_1 &= c + dR_0 + (\delta - k\hat{T}) = c + dR_0 + \frac{N\delta^2\beta}{N\delta\beta + kc} > 0, \\ a_2 &= dcR_0 + d\delta(R_0 - 1) + d(\delta - k\hat{T}) = dcR_0 + d\delta(R_0 - 1) + \frac{N\delta^2\beta}{N\delta\beta + kc}, \\ a_3 &= dc\delta(R_0 - 1). \end{aligned} \tag{C.3}$$

It's clear that Eq (C.2) always has a negative root $\lambda = -\delta$ with multiplicity n. When $R_0 > 1$, the inequalities $a_2 > 0$ and $a_3 > 0$ hold. Moreover, we have

$$\begin{aligned} a_1a_2 - a_3 &= \Big(c + dR_0 + (\delta - k\hat{T})\Big)\Big(dcR_0 + d\delta(R_0 - 1) + d(\delta - k\hat{T})\Big) - dc\delta(R_0 - 1) \\ &= c\Big(dcR_0 + d(\delta - k\hat{T})\Big) + \Big(dR_0 + (\delta - k\hat{T})\Big)\Big(dcR_0 + d\delta(R_0 - 1) + d(\delta - k\hat{T})\Big) \\ &> 0. \end{aligned}$$

It follows the Routh-Hurwitz criterion that all the roots of (C.2) have negative real parts. Thus, the infected equilibrium E_1 is locally asymptotically stable when it exists, i.e., when $R_0 > 1$. This completes the proof.

AIMS Press