Journal of Theoretical Biology 509 (2021) 110502

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

# Modeling HIV multiple infection

Ting Guo [a,c], Zhipeng Qiu [a], Kosaku Kitagawa [b], Shingo Iwami [b], Libin Rong [c,*]

[a] School of Science, Nanjing University of Science and Technology, Nanjing 210094, China
[b] Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka 8190395, Japan
[c] Department of Mathematics, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

*Article history:*
Received 1 May 2020
Revised 9 August 2020
Accepted 19 September 2020
Available online 28 September 2020

*Keywords:*
HIV infection
Mathematical model
Multiple infections
Cell-to-cell transmission
Sequential cell-free virus infection

ABSTRACT

Multiple infection of target cells by human immunodeficiency virus (HIV) may lead to viral escape from host immune responses and drug resistance to antiretroviral therapy, bringing more challenges to the control of infection. The mechanisms underlying HIV multiple infection and their relative contributions are not fully understood. In this paper, we develop and analyze a mathematical model that includes sequential cell-free virus infection (i.e. one virus is transmitted each time in a sequential infection of target cells by virus) and cell-to-cell transmission (i.e. multiple viral genomes are transmitted simultaneously from infected to uninfected cells). By comparing model prediction with the distribution data of proviral genomes in HIV-infected spleen cells, we find that multiple infection can be well explained when the two modes of viral transmission are both included. Numerical simulation using the parameter estimates from data fitting shows that the majority of T cell infections are attributed to cell-to-cell transmission and this transmission mode also accounts for more than half of cell's multiple infections. These results suggest that cell-to-cell transmission plays a critical role in forming HIV multiple infection and thus has important implications for HIV evolution and pathogenesis.



## 1. Introduction

HIV infection is still a major global public health issue. Mathematical modeling has proven to be a powerful tool to study viral pathogenesis and provided valuable insights into the viral infection dynamics (Zitzmann and Kaderali, 2018; Kumberger et al., 2018). Over the past decades, there has been a great effort in the mathematical modeling of HIV infection (see Perelson and Ribeiro, 2013; Hill et al., 2018; Perelson and Nelson, 1999; Rong and Perelson, 2009; Stafford et al., 2000; Kumberger et al., 2016; Lai and Zou, 2014 and references therein). These models describe HIV replication dynamics and the effects of antiretroviral therapy. Most models only considered cell-free HIV infection, i.e., target cell is infected by cell-free virus upon contact. However, multiple infections of target cells have been observed in experiments. For example, Jung et al. (2002) found that HIV-infected CD4+ T cells from the spleens of two patients carry up to 8 proviruses, with a mean of 3 to 4 proviruses per cell. A growing body of studies indicated that HIV multiple infection occurs far more frequently than single infection (Levy et al., 2004; Dang et al., 2004; Chen et al., 2005). A question thus arises as to what mechanism is responsible for the high-frequency of HIV multiple infection.

A possible explanation for HIV multiple infection is cell-to-cell transmission (see Fig. 1(a)). Multiple viral genomes are simultaneously transmitted in individual infectious events when infected cells encounter uninfected cells and form viral synapses (Jolly and Sattentau, 2004; Agosto et al., 2015; Portillo et al., 2011; Mothes et al., 2010; Sigal et al., 2011). Chen et al. (2005) showed that multiple virions can be transmitted from dendritic cells to T cells during cell-mediated HIV transmission. More studies further supported cell-to-cell transmission and showed that an infected cell can transfer a various number of virions directly to the target cells (Chen et al., 2007; Hübner et al., 2009; Komarova and Wodar, 2013; Graw and Perelson, 2016).

Dixit and Perelson (2005) studied another multiple infection-involved mechanism in which target cells are sequentially infected by cell-free virions and each infectious contact results in the transmission of one viral genome, i.e. sequential cell-free virus infection (see Fig. 1(b)). In their study, the number of virions was assumed to be several orders of magnitude higher than the number of target cells, which was the feature of many *in vitro* experiments (Dixit and Perelson, 2005). Thus, cell-to-cell transmission may be neglected and cell-free virus infection might be the major contributor to multiple infections (Dimitrov et al., 1993). Such multiple infections have been observed in some other *in vitro* experiments (Levy et al., 2004; Dang et al., 2004; Ito et al., 2017, 2018).

* Corresponding author.
*E-mail address:* libinrong@ufl.edu (L. Rong).

https://doi.org/10.1016/j.jtbi.2020.110502
0022-5193/

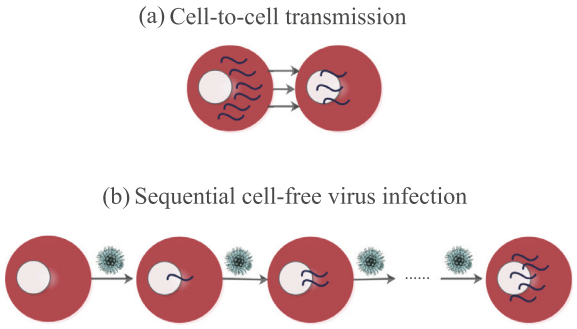


**Fig. 1.** Schematic illustration of two possible mechanisms for HIV multiple infections. (a) Cell-to-cell transmission: multiple viral genomes are simultaneously transmitted from an infected cell to an uninfected target cell through the formation of virological synapse. (b) Sequential cell-free virus infection: target cells have sequential contact with cell-free virions and one viral genome is transmitted in each contact.

It is important to obtain quantitative insights into multiple infections using mathematical models. Some studies have provided critical first steps in this regard. For example, Ito et al. (2017) reproduced the datasets in cell-free HIV single and double infection experiments using a mathematical model, in which the variation in susceptibility of target cells was assumed to be a continuous distribution. In a subsequent paper (Ito et al., 2018), they showed that dividing the cell populations into two susceptible subpopulations could explain their HIV coinfection experimental data, while increasing the number of subpopulations did not improve the fitting. Dixit and Perelson (2005) proposed an HIV dynamical model of multiple infections and reproduced the experimentally observed scaling law under certain conditions. These models assumed that multiple infections are the result of cell-free virus infection. In Dixit and Perelson (2004), they constructed two models and showed that cell-to-cell transmission and sequential cell-free virus infection can respectively explain the experimental data from Jung et al. (2002). Komarova and Wodar (2013) formulated a class of models that include both cell-to-cell transmission and sequential cell-free virus infection. The models were used to study how the number of viruses passed per synapse affects the basic reproduction number and evolutionary dynamics of two viral strains.

In this paper, we develop a general mathematical model that includes two virus transmission modes and distinguishes the number of proviral genomes in multiple infections. We analyze the model by deriving the basic reproductive number and studying the stability of the steady states. Using the infected steady states, we calculate the distribution of infected cells with different numbers of proviruses. To explain how multiple infections are orchestrated, we fit models with different assumptions of the mechanisms for multiple infections to the distribution data of proviruses from spleen cells of two patients Jung et al., 2002. Using the model with best-fitted parameter values, we evaluate the relative contributions to CD4+ T cell infection from the two routes of virus spread, as well as two classes of infected cells (i.e. infected cells with single and multiple proviruses). On this basis, the relative contributions of two virus infection modes to two classes of infected cells can also be quantitatively investigated.

## 2. Model formulation and analysis

### 2.1. A model with sequential virus infection and cell-to-cell transmission

In this section, we formulate a general ordinary differential equation model with HIV multiple infections from sequential cell-free virus infection and cell-to-cell transmission. Assume that a target cell can be infected by up to $n$ viruses. Let $T$ and $V$ be the concentration of uninfected CD4+ T cells and virus, respectively, and $I_i$ be the concentration of CD4+ T cells infected by $i$ virions, $1 \leqslant i \leqslant n$, where the number $i$ is the cell's multiplicity of infection (MOI). Based on standard viral dynamical models (Perelson and Nelson, 1999; Nowak and May, 2000), we develop a model with HIV multiple infections, described by the following system

$$\begin{cases} \frac{dT}{dt} = \Lambda - \mu T(t) - \beta_1 T(t)V(t) - \kappa T(t)\sum_{j=1}^{n} I_j(t), \\ \frac{dI_1}{dt} = \beta_1 T(t)V(t) - \beta_2 I_1(t)V(t) + f_1\kappa T(t)\sum_{j=1}^{n} I_j(t) - \delta I_1(t), \\ \frac{dI_i}{dt} = \beta_i I_{i-1}(t)V(t) - \beta_{i+1} I_i(t)V(t) + f_i\kappa T(t)\sum_{j=1}^{n} I_j(t) - \delta I_i(t), \\ \quad i = 2, 3, \ldots, n-1, \\ \frac{dI_n}{dt} = \beta_n I_{n-1}(t)V(t) + f_n\kappa T(t)\sum_{j=1}^{n} I_j(t) - \delta I_n(t), \\ \frac{dV}{dt} = p\sum_{i=1}^{n} I_i(t) - cV(t). \end{cases} \tag{1}$$

In the model, uninfected cells are assumed to be produced at a rate $\Lambda$, die with the first order rate constant $\mu$. Parameter $\beta_1$ is the infection rate of uninfected cells by free virus and $\beta_i$ $(i = 2, 3, \ldots, n)$ represents the reinfection rate of infected cells with $i-1$ MOI by free virus. Constant $\kappa$ is the rate of cell-to-cell viral transmission. The parameter $f_i$ $(i = 1, 2, \ldots, n)$ denotes the probability of transmitting $i$ viral genomes during cell-to-cell transmission. Thus, we have $\sum_{i=1}^{n} f_i = 1$. Infected cells is assumed to die at the same rate $\delta$ although we can make it different for infected cells with different numbers of proviral genomes. We have $\delta \geqslant \mu$ because of the cytopathic effect of viral proteins and cytotoxic T cell killing. Free virions are assumed to be produced by infected cells at a rate $p$ and are cleared at a rate $c$. All the parameters are assumed to be positive constants.

### 2.2. Model analysis

In this section, we derive the basic reproduction number and investigate the existence and stability of equilibria. Any equilibrium of (1) must satisfy the following equalities:

$$\begin{cases} \Lambda - \mu T - \beta_1 TV - \kappa T\sum_{j=1}^{n} I_j = 0, \\ \beta_1 TV - \beta_2 I_1 V + f_1\kappa T\sum_{j=1}^{n} I_j - \delta I_1 = 0, \\ \beta_i I_{i-1}V - \beta_{i+1} I_i V + f_i\kappa T\sum_{j=1}^{n} I_j - \delta I_i = 0, \quad i = 2, 3, \ldots, n-1 \\ \beta_n I_{n-1}V + f_n\kappa T\sum_{j=1}^{n} I_j - \delta I_n = 0, \\ p\sum_{i=1}^{n} I_i - cV = 0. \end{cases} \tag{2}$$

Obviously, the system (1) always has an infection-free equilibrium $E_0 = (\Lambda/\mu, 0, 0, \ldots, 0, 0)$.The stability of $E_0$ can be determined by the next generation operator method. Following the notations in den Driessche and Watmough (2002), the matrices $\mathcal{F}$ and $\mathcal{V}$, i.e. the new infection and transfer matrices, are respectively given by

$$\mathcal{F}=\begin{pmatrix} \kappa f_1\frac{\Lambda}{\mu} & \kappa f_1\frac{\Lambda}{\mu} & \cdots & \kappa f_1\frac{\Lambda}{\mu} & \beta_1\frac{\Lambda}{\mu} \\ \kappa f_2\frac{\Lambda}{\mu} & \kappa f_2\frac{\Lambda}{\mu} & \cdots & \kappa f_2\frac{\Lambda}{\mu} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \kappa f_n\frac{\Lambda}{\mu} & \kappa f_n\frac{\Lambda}{\mu} & \cdots & \kappa f_n\frac{\Lambda}{\mu} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}, \quad \mathcal{V}=\begin{pmatrix} \delta & 0 & \cdots & 0 & 0 \\ 0 & \delta & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \delta & 0 \\ -p & -p & \cdots & -p & c \end{pmatrix}.$$

Thus, the next generation matrix for HIV infection, denoted by $\mathcal{M}$, can be calculated by

$$\mathcal{M}=\mathcal{F}\mathcal{V}^{-1}=\begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathbf{0} & 0 \end{pmatrix}, \tag{3}$$

where

$$\mathcal{M}_{11}=\begin{pmatrix} f_1\frac{\kappa}{\delta}\frac{\Lambda}{\mu}+\frac{p}{\delta}\frac{\beta_1}{c}\frac{\Lambda}{\mu} & f_1\frac{\kappa}{\delta}\frac{\Lambda}{\mu}+\frac{p}{\delta}\frac{\beta_1}{c}\frac{\Lambda}{\mu} & \cdots & f_1\frac{\kappa}{\delta}\frac{\Lambda}{\mu}+\frac{p}{\delta}\frac{\beta_1}{c}\frac{\Lambda}{\mu} \\ f_2\frac{\kappa}{\delta}\frac{\Lambda}{\mu} & f_2\frac{\kappa}{\delta}\frac{\Lambda}{\mu} & \cdots & f_2\frac{\kappa}{\delta}\frac{\Lambda}{\mu} \\ \vdots & \vdots & \ddots & \vdots \\ f_n\frac{\kappa}{\delta}\frac{\Lambda}{\mu} & f_n\frac{\kappa}{\delta}\frac{\Lambda}{\mu} & \cdots & f_n\frac{\kappa}{\delta}\frac{\Lambda}{\mu} \end{pmatrix}_{n\times n},$$

and

$$\mathcal{M}_{12}=\left(\left(\frac{\beta_1}{c}\frac{\Lambda}{\mu} \quad 0 \quad \cdots \quad 0\right)_{1\times n}\right)^T.$$

Note that $\mathrm{rank}(\mathcal{M})=\mathrm{rank}(\mathcal{M}_{11})=1$. It follows from the properties of eigenvalues that the matrix $\mathcal{M}$ has only one nonzero eigenvalue $tr(\mathcal{M})$, where $tr(\mathcal{M})$ is the trace of $\mathcal{M}$. Thus, the basic reproduction number can be calculated by

$$\mathcal{R}_0=\rho(\mathcal{M})=tr(\mathcal{M})=\frac{\kappa}{\delta}\frac{\Lambda}{\mu}+\frac{p}{\delta}\frac{\beta_1}{c}\frac{\Lambda}{\mu},$$

where $\rho(\mathcal{M})$ represents the spectral radius of matrix $\mathcal{M}$. It represents the average number of new cell infections induced by introduction of a single infected cell into a wholly susceptible target cell population. The first term is the contribution from cell-to-cell transmission and the second term is the contribution from cell-free virus infection.

Next, we study the global dynamics of the system (1). We first show the existence of the positive equilibrium of system (1). From the last equation in Eq. (2), we have

$$\sum_{i=1}^{n} I_i=\frac{c}{p}V. \tag{4}$$

Adding the first $n+1$ equations of (2), and the equations from the second to the $(n+1)$th of (2), we can obtain

$$T=\frac{\Lambda}{\mu}-\frac{\delta}{\mu}\frac{c}{p}V, \tag{5}$$

and

$$\beta_1 TV+\kappa T\sum_{j=1}^{n} I_j=\delta\sum_{j=1}^{n} I_j. \tag{6}$$

Substituting (4) and (5) into (6) yields that

$$\beta_1\left(\frac{\Lambda}{\mu}-\frac{\delta c}{\mu p}V\right)V+\kappa\left(\frac{\Lambda}{\mu}-\frac{\delta c}{\mu p}V\right)\frac{c}{p}V=\delta\frac{c}{p}V. \tag{7}$$

Eq. (7) is a quadratic equation of $V$. Clearly, $V=0$ solves (7). Substituting $V=0$ into (2), we obtain that all state variables, except for $T$, are zero and $T$ equals $\Lambda/\mu$. This again shows that the infection-free equilibrium $E_0=(\Lambda/\mu,0,0,\ldots,0,0)$ always exists. When $V\neq 0$, we cancel $V$ from both sides of (7) and obtain

$$\beta_1\left(\frac{\Lambda}{\mu}-\frac{\delta c}{\mu p}V\right)+\kappa\left(\frac{\Lambda}{\mu}-\frac{\delta c}{\mu p}V\right)\frac{c}{p}=\delta\frac{c}{p}. \tag{8}$$

Solving (8), we have

$$V=\frac{\mu p\left[\frac{\kappa}{\delta}\frac{\Lambda}{\mu}+\frac{p}{\delta}\frac{\beta_1}{c}\frac{\Lambda}{\mu}-1\right]}{\beta_1 p+\kappa c}=\frac{\mu p[\mathcal{R}_0-1]}{\beta_1 p+\kappa c}:=V^*,$$

which is the unique positive root of (8) if and only if $\mathcal{R}_0>1$. Substituting $V=V^*$ into (5), we have

$$T=\frac{\Lambda}{\mu}-\frac{\delta}{\mu}\frac{c}{p}V^*=\frac{\delta c}{\beta_1 p+\kappa c}:=T^*>0.$$

Substituting $V=V^*$ and $T=T^*$ into the equations from the second to the $(n+1)$th of (2) and solving the equations for $I_1,I_2,\ldots,I_n$, we obtain

$$\begin{aligned} I_1&=\frac{\beta_1 p+f_1\kappa c}{\delta p+\beta_2 pV^*}T^*V^*:=I_1^*,\\ I_i&=\frac{\beta_i pI_{i-1}^*V^*+f_i\kappa cT^*V^*}{\delta p+\beta_{i+1}pV^*}:=I_i^*,\quad i=2,3,\ldots,n-1,\\ I_n&=\frac{\beta_n pI_{n-1}^*V^*+f_n\kappa cT^*V^*}{\delta p}:=I_n^*. \end{aligned} \tag{9}$$

From the above expressions, we have $I_i^*>0$ $(i=1,2,\ldots,n)$ if and only if $V^*>0$. Thus, the system (1) has a unique positive equilibrium if and only if $\mathcal{R}_0>1$.

The results on the existence of the positive equilibrium can be summarized in the following theorem.

**Theorem 2.1.** *The system (1) has no positive equilibrium if $\mathcal{R}_0\leqslant 1$, and has a unique positive equilibrium $E^*$ if $\mathcal{R}_0>1$.*

*Furthermore, the global dynamics of system (1) can be stated as follows. The proof is given in the Appendix.*

**Theorem 2.2.** *If $\mathcal{R}_0\leqslant 1$, then the infection-free equilibrium $E_0$ of system (1) is globally asymptotically stable; if $\mathcal{R}_0>1$, then the positive equilibrium $E^*$ is globally asymptotically stable.*

## 3. Data analysis and numerical results

Jung et al. (2002) found that the splenocytes in HIV-infected patients harbor multiple copies of proviruses. Multiple viral genomes may be obtained by cell-to-cell transmission, sequential cell-free virus infection or a combination of the two routes. In this section, we consider these three cases to test which mechanism can best explain the distribution of proviruses in the patients (Jung et al., 2002).

We define $\psi_i$ $(i=1,2,\ldots,n)$ as the proportion of infected cells with $i$ proviruses in the total infected cells at steady state $E^*$, that is,

$$\psi_i=I_i^*/\left(\sum_{i=1}^{n} I_i^*\right).$$

By the least square method, we fit $\psi_i$ to the experimental data from Jung et al. (2002) under different assumptions. To compare the best fits, the sum of squared residuals (SSR) and the Akaike information criterion (AIC) are calculated. Using model (1) with the best-fitted parameter values, we evaluate the relative contributions to CD4+ T cell infection from two types of infection modes (i.e. cell-to-cell transmission and sequential cell-free virus infection) and two classes of infected cells (i.e. infected cells with single and multiple proviruses). The relative contributions to two classes of infected cells from two types of infection modes will also be determined.

### 3.1. Comparison with patient data

Using the fluorescence *in situ* hybridization, Jung et al. measured the proviral copy number from infected splenocytes in two patients (Jung et al., 2002). Although two patients have different clinical profiles and viral loads, the frequency distributions of proviral copy number are remarkably similar (see Fig. 1b of Jung et al., 2002). The number of proviruses ranged from one to eight per cell, with an average of about 3.2. Thus, we simulate the model with a maximum of 8 proviruses per cell (i.e. $n = 8$). Dividing the number of cells containing a given number of proviruses by the total number of cells from each patient (113 cells for patient R, 103 cells for patient B), we compare with model prediction of the frequency distribution.

We fix some parameter values based on modeling literature and experimental data (Mohri et al., 1998; Bofill et al., 1992; Rong and Perelson, 2009; Guo and Qiu, 2019; Wang and Rong, 2019; Wang et al., 2017). The death rate of uninfected cells $\mu$ is assumed to be 0.01 $day^{-1}$ (Mohri et al., 1998; Wang and Rong, 2019). Because an uninfected individual has a CD4+ T cell count of approximately $10^6$ per $ml$ blood (Bofill et al., 1992), we obtain that the generation rate of uninfected cells $\Lambda$ is $10^6 \times 0.01 = 10^4\ ml^{-1}day^{-1}$ from the steady state of target cells before infection. The death rate of infected cells $\delta$ and the clearance rate of virus $c$ are 1 $day^{-1}$ and 23 $day^{-1}$ (Rong and Perelson, 2009; Guo and Qiu, 2019; Wang and Rong, 2019; Wang et al., 2017), respectively. We fit the model to experimental data and estimate the remaining parameters, including the infection rate of infected cells with MOI $= i - 1$ by cell-free virus, i.e. $\beta_i$ $(i = 1, 2, \ldots, n)$, the rate of cell-to-cell viral transmission $\kappa$, the probability of transmitting $i$ viral genomes during cell-to-cell transmission $f_i$ $(i = 1, 2, \ldots, n)$, and the viral production rate $p$. Based on the mechanisms that may generate multiple infections of cells, we consider the following three cases.

**Case A: Only sequential cell-free virus infection**

If there is only sequential cell-free virus infection, the rate of cell-to-cell viral transmission $\kappa$ is 0. In this case, we fit $\psi_i$ $(i = 1, 2, \ldots, 8)$ to the data (Jung et al., 2002). Because of the limited experimental data, we further make assumption on the infection rate $\beta_i$, $(i = 1, 2, \ldots, 8)$. Komarova et al. (2013) assumed that $\beta_i$ is independent of $i$, i.e., $\beta_i$ is equal to a constant $\beta$. On the other hand, due to the virus-induced CD4 down-modulation (Levy et al., 2004), infected cells with a higher MOI are less likely to be further infected by free virus during sequential cell-free virus infection. Thus, $\beta_i$ can be a decreasing function with respect to the cell's MOI (i.e. parameter $i$). To evaluate the effect of these different assumptions on data fitting, we consider the following two cases.

If the infection rate by cell-free virus does not depend on the MOI (i.e. $\beta_i = \beta$), then the best fits are shown in Fig. 2(a,b) for two patients (black solid line with asterisk). On the basis of the best fits, we estimated $\beta = 2.97 \times 10^{-8}\ ml/day$ and $p = 1928.7\ cell^{-1}day^{-1}$. Two model predictions are similar in the distribution of infected cells. There are only singly infected cells and no multiply infected cells (Fig. 2(a,b)). This does not agree with the observation that each infected cell has an average of three or four proviruses (Jung et al., 2002). Thus, the model with only sequential cell-free virus infection and constant infection rate (i.e. $\beta_i$ is the same constant for all $i$) yields a very poor fit to patient data.

Next, we consider a decreasing infection rate as the number of proviruses increases. For simplicity, we use the decreasing function $\beta_i = \beta_1 \times r^{(i-1)}$ $(i = 2, 3, \ldots, 8)$ as the reinfection rate of infected CD4+ T cells with $i - 1$ MOI, where $0 < r < 1$ is constant that characterizes how fast the infection rate declines as the MOI increases. Another model assumed that the infection rate is exponentially declining but introduced additional parameters (Dixit and Perelson, 2005). A simplification of this model, allowing up to two infections per cell, showed that $\beta_2 = 0.7 \times \beta_1$ can quantitatively capture the observations of the frequencies of coinfection *in vitro* (Suryavanshi and Dixit, 2007). Thus, we assume the decay constant to be $r = 0.7$ such that $\beta_2 = 0.7 \times \beta_1$. The best fits are shown in Fig. 2(c,d). Consistent with the results obtained with $\beta_i = \beta$, the model predictions are unable to capture the observed patient data (Jung et al., 2002). We also examine whether the variation of the decay constant $r$ can improve the fitting. Using other values of $r$ during fitting, we find that the model predictions are not sensitive to the variations in $r$ (figures not shown). Therefore, the model with a decreasing infection rate $\beta_i$ again provides a poor fit to the data in two patients (Jung et al., 2002).

In summary, the model (1) that includes only cell-free virus infection without cell-to-cell transmission fails to reproduce the patient data (Jung et al., 2002), no matter if the infection rate of free virus is independent or decreasing of MOI. In a previous study (Dixit and Perelson, 2004), Dixit and Perelson showed cell-to-cell transmission may explain the observation in Jung et al. (2002). Although their model is different from the one in our paper, their result inspires us to consider cell-to-cell virus transmission below.

**Case B: Only cell-to-cell transmission**

If only cell-to-cell viral transmission results in more than one provirus per spleen cell, then the infection rate of free virus $\beta_i$ is 0 for $i = 1, 2, \ldots$. Dixit and Perelson (2004) supposed that the number of genomes transmitted per infectious cell-associated contact obeys the modified Poisson distribution, i.e. $f_i = \frac{e^{-\gamma}}{1-e^{-\gamma}} \frac{\gamma^i}{i!}$, where $\gamma$ represents the average number of genomes transmitted per infectious cell-to-cell transmission. We also use this distribution as an example to describe cell-to-cell transmission. We fit the proportion $\psi_i$ $(i = 1, 2, \ldots, 8)$ to the distribution data of infected cells harboring different proviral copy numbers in each patient. The best fits, shown in Fig. 3(a,b), indicate that model (1) with cell-to-cell transmission alone provides a reasonable fit to the infection cell distribution data.

Although the model with modified Poisson distribution of $f_i$ agrees with the data in the two patients (Jung et al., 2002), the sensitivity of this result to the change in the distribution of $f_i$ remains unclear. We examine whether other distributions of $f_i$ alter the fitting result. We choose binomial distribution as an example. During cell-to-cell transmission, there are two possible outcomes: success or failure. Thus, the transfer of a viral genome can be considered as a Bernoulli trial. The result of one trial is independent of the others. We can assume that each viral genome has a probability $\theta$ to establish a successful transfer, or fail with the probability $1 - \theta$. Thus, the binomial distribution $f_i = \binom{n}{i}\theta^i(1-\theta)^{n-i}$ can be used to describe the probability of $i$ successes during the transmission of $n$ viral genomes, where $n = 8$ is the maximum number of viral genomes observed in the patient data from Jung et al. (2002). The binomial distribution has also been used in Refs. Komarova and Wodar (2013) and Komarova et al. (2013), although they introduced the cell-to-cell transmission in a different way from this study. Using the same fitting approach as described above, we obtain the best fits in Fig. 3(c,d), which also generate good reproduction of the patient data. Taken together, these results indicate that the model with cell-to-cell transmission alone can provide good fits to the patient data when $f_i$ follows the modified Poisson distribution or binomial distribution (see Fig. 3). However, it remains unclear which distribution generates a better model prediction. To compare the best fits with these two distributions, we calculate the values of SSR and AIC. These values evaluate the goodness of different model fits. More specifically, lower SSR indicates better fitting and lower AIC value indicates a better model. The SSR and AIC are given by

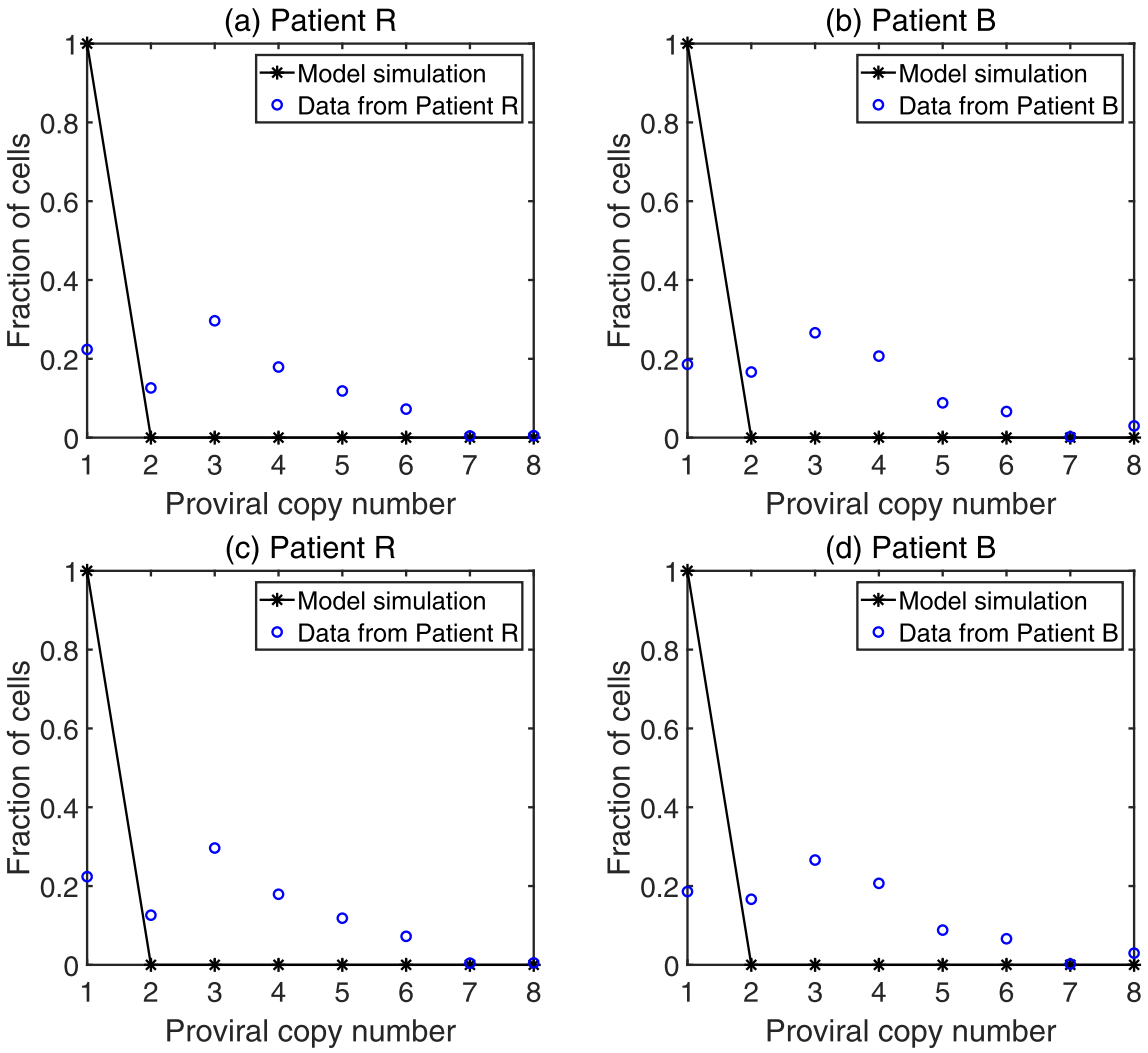


**Fig. 2.** Data from two patients (blue circles) compared with model predictions (black solid lines with asterisks). The fitting assumes that only sequential cell-free virus infection exists, i.e. $\kappa = 0$. Upper panel: The infection rate by free virus is assumed to not depend on MOI (i.e. $\beta_i = \beta$). Low panel: The infection rate by free virus $\beta_i$ is assumed to be a decreasing function of cell's MOI. All model predictions present the same results: no multiply infected cells. The patient data are the ratios of the number of cells containing $i$ proviruses ($i = 1, 2, \ldots, 8$) to the total number of infected cells. The other parameter values used in the simulations are: $\mu = 0.01\ day^{-1}, \Lambda = 10^4\ ml^{-1}day^{-1}, \delta = 1\ day^{-1}$ and $c = 23\ day^{-1}$.

$$\mathrm{SSR} = \sum_{i=1}^{n}(x_i - \tilde{x}_i)^2, \qquad \mathrm{AIC} = n\ln(\mathrm{SSR}/n) + 2p, \quad i = 1, 2, \ldots n, \tag{10}$$

where $x_i$ is the predicted value of $\psi_i$ calculated from model (1) with $\beta_i = 0$ and different distributions of $f_i$, $\tilde{x}_i$ is the proportion of infected cells harboring $i$ proviral copies, $n = 8$ is the number of data points, and $p$ is the number of parameters. For each patient, the values of SSR and AIC based on the best fits are listed in Table 1. The model (1) with only cell-to-cell transmission and the modified Poisson distribution of $f_i$, denoted by $model_{cm}$ in Table 1, has lower values of SSR and AIC than $model_{cb}$, the model (1) with only cell-to-cell transmission and binomial distribution of $f_i$. Thus, when $f_i$ follows a modified Poisson distribution, the model (1) with cell-to-cell transmission yields a slightly better fit but the improvement is not significant.

In addition to the modified Poisson distribution and binomial distribution, other distributions such as the negative binomial distribution can also be used in the fitting. However, the model with cell-to-cell transmission alone does not capture the bimodal distribution in the patient data (Jung et al., 2002) with these distributions. The distribution of the proviral copy number in infected splenocytes has two peaks at $i = 1$ and $i = 3$, whereas the fits of model (1) with only cell-to-cell transmission only displays a single peak at $i = 3$ (see Fig. 3). Recall that the fits of model (1) with only sequential cell-free virus infection have a single peak at $i = 1$ (see Fig. 2). We speculate that including two modes of viral spread may explain the bimodal distribution of patient data. To test this hypothesis, we conduct the fitting using a model with both transmission routes.

**Case C: Both cell-free virus infection and cell-to-cell transmission**

Results in Table 1 suggest that the model (1) with cell-to-cell transmission alone and the modified Poisson distribution of $f_i$ has provided a better fit. However, this does not mean that model (1) with two transmission routes and the modified Poisson distribution of $f_i$ provides the best fit. This is because cells infected via the cell-free mode can transfer viral genomes to uninfected cells via cell-to-cell transmission, which certainly affects the frequency of infected cells with various proviral copies and the result of data fitting. Thus, we will discuss the following four cases on the basis of the assumptions on $\beta_i$ and $f_i$.

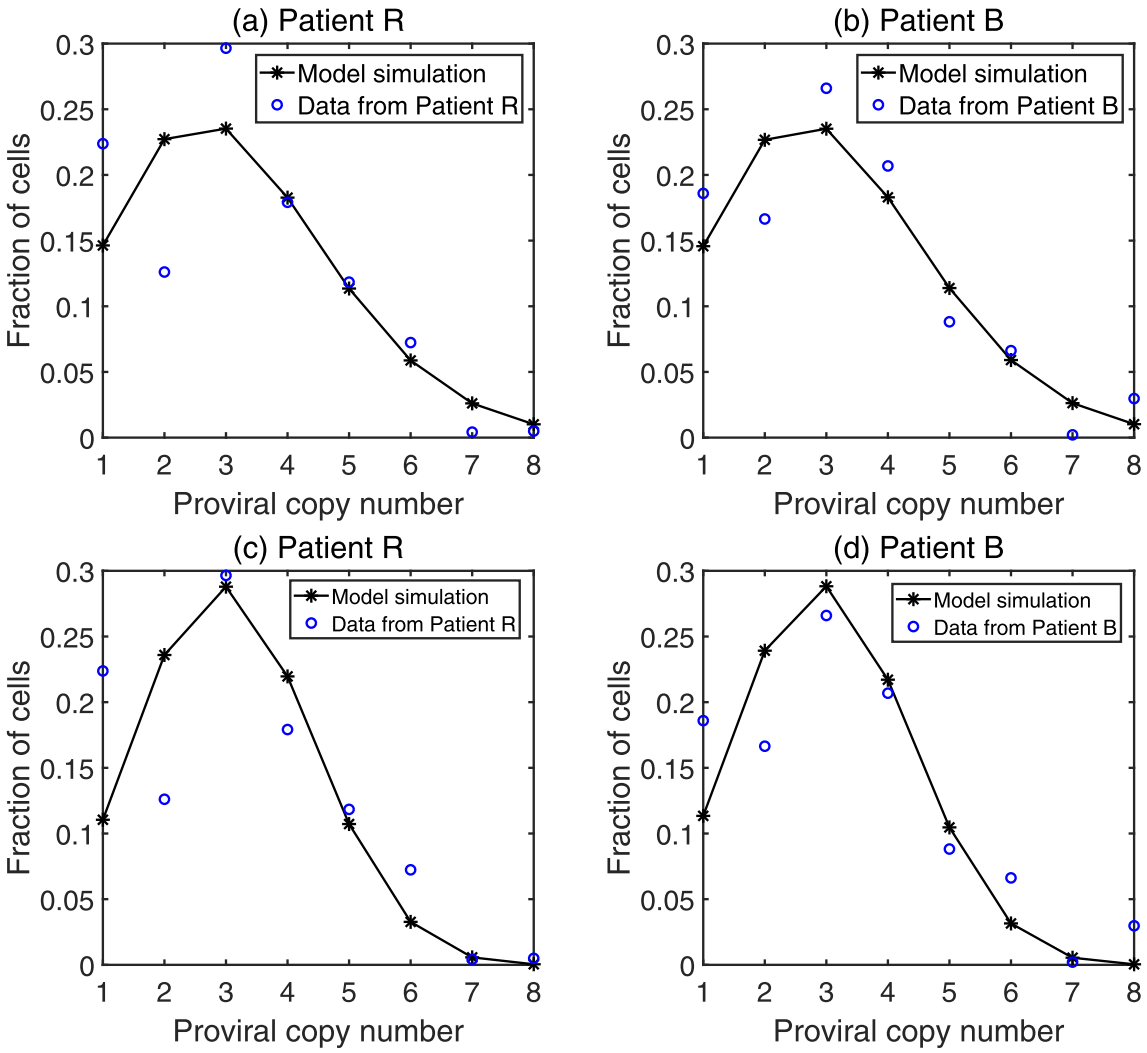


**Fig. 3.** The best fits of the predicted proportion $\psi_i$ (black solid lines with asterisks) to the distribution data (blue circles) from two patients. The fitting assumes that only cell-to-cell transmission exists, i.e. $\beta_i = 0$ ($i = 1, 2, \ldots, 8$). Upper panel: The number of genomes transmitted per infectious cell-to-cell contact is assumed to obey the modified Poisson distribution. Low panel: We assume that $f_i$ obeys binomial distribution. Model fitting with cell-to-cell transmission alone provides good fit to the data from two patients, and shows a peak of the proportion at $i = 3$. The other parameter values are the same as in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**
Comparison of the best fits using different models.*

| | SSR | | | | AIC | | | |
|---|---|---|---|---|---|---|---|---|
| Patient | $Model_{cm}$ | $Model_{cb}$ | $Model_{bcm}$ or $Model_{bdm}$ | $Model_{bcb}$ or $Model_{bdb}$ | $Model_{cm}$ | $Model_{cb}$ | $Model_{bcm}$ or $Model_{bdm}$ | $Model_{bcb}$ or $Model_{bdb}$ |
| R | 0.021 | 0.028 | 0.012 | 0.007 | −41.657 | −39.155 | −44.01 | −48.3 |
| B | 0.008 | 0.014 | 0.006 | 0.003 | −48.872 | −45.076 | −49.56 | −55.1 |

$Model_{cm}$ is model (1) with only cell-to-cell transmission and modified Poisson distribution for $f_i$;
$Model_{cb}$ is model (1) with only cell-to-cell transmission and binomial distribution for $f_i$;
$Model_{bcm}$ is model (1) with two transmission routes, constant $\beta_i$, and modified Poisson distribution for $f_i$;
$Model_{bcb}$ is model (1) with two transmission routes, constant $\beta_i$, and binomial distribution for $f_i$;
$Model_{bdm}$ is model (1) with two transmission routes, decreasing $\beta_i$, and modified Poisson distribution for $f_i$;
$Model_{bdb}$ is model (1) with two transmission routes, decreasing $\beta_i$, and binomial distribution for $f_i$.
* The values of SSR and AIC were calculated by Eq. (10).

When $\beta_i$ ($i = 1, 2, \ldots, 8$) is the same constant and $f_i$ follows the modified Poisson distribution, we fit the proportion $\psi_i$ to the patient data from Jung et al. (2002). As shown in Fig. 4(a,b), the best fits predict a bimodal distribution, with two peaks at $i = 1$ and $i = 3$. For the case that $f_i$ obeys binomial distribution (assuming $\beta_i$ is still independent of $i$), we obtain the same results, as shown in Fig. 4(c,d). To determine which one of two distributions is better for data fitting, we calculate the values of SSR and AIC in different cases for two patients. These values are listed in Table 1. We find that the model (1) with the binomial distribution of $f_i$ (i.e. $Model_{bcb}$ of Table 1) has smaller values of SSR and AIC than model (1) with the modified Poisson distribution of $f_i$ (i.e. $Model_{bcm}$ of Table 1), suggesting that the former gives a better fit to the patient data.

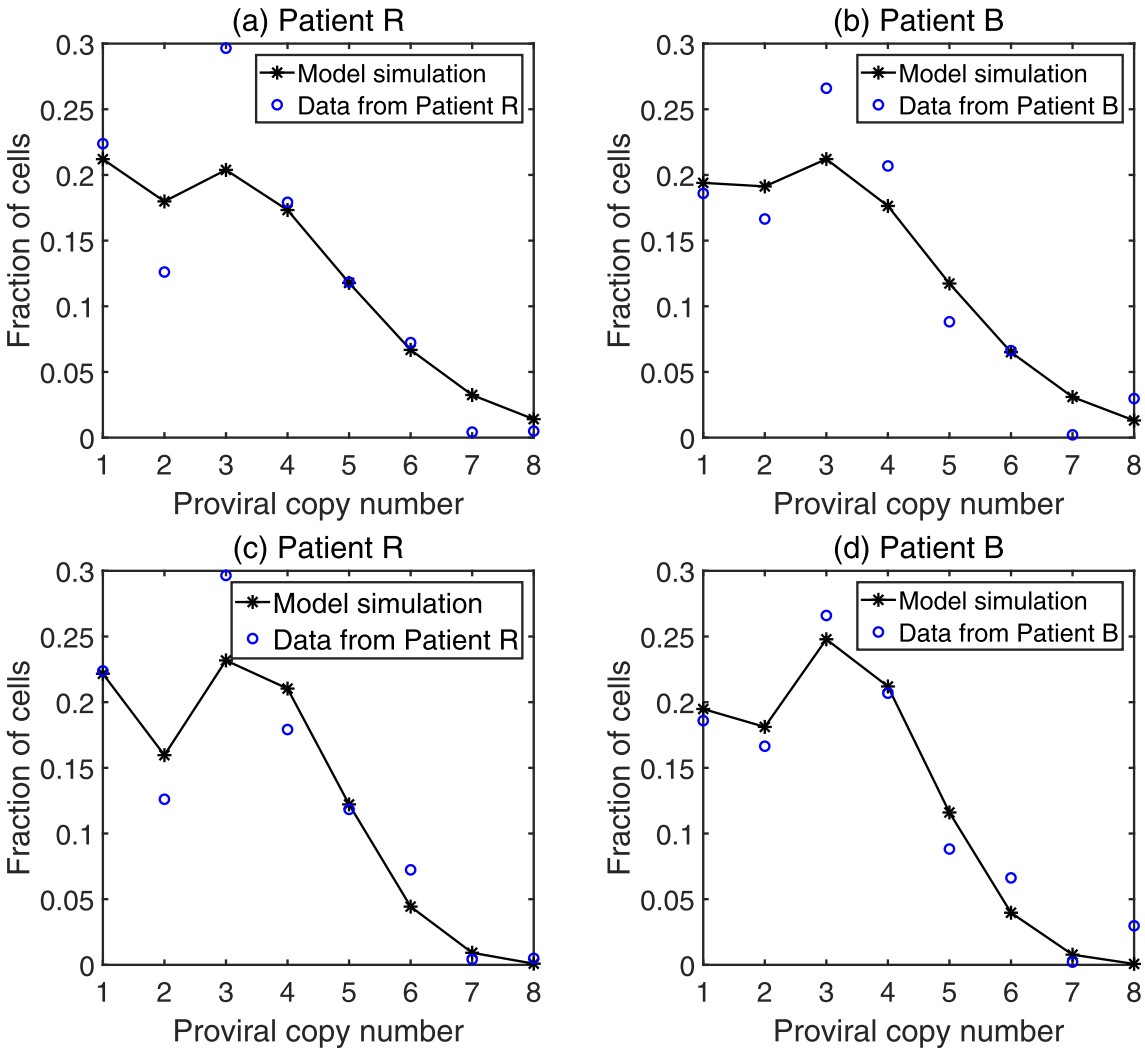


**Fig. 4.** Data fitting using the model with both sequential cell-free virus infection and cell-to-cell transmission. The infection rate $\beta_i$ is assumed to be independent of the MOI. Upper panel: The number of genomes transmitted per infectious cell-to-cell contact obeys the modified Poisson distribution. Low panel: We assume that $f_i$ obeys binomial distribution. The fitting captures the bimodal distribution of data with peaks at $i = 1$ and $i = 3$. The other parameter values are the same as in Fig. 2.

Assuming $\beta_i$ is a decreasing function of $i$ (we choose $\beta_i = \beta_1 \times 0.7^{i-1}$ as an example in the fitting), the fitting results with different distributions of $f_i$ are presented in Fig. 5. Consistent with our above findings, the model (1) with decreasing $\beta_i$ captures the bimodal distribution of the patient data, providing a good fit. The values of SSR and AIC for the different distributions of $f_i$ are shown in Table 1. Interestingly, given a specific distribution of $f_i$, the same values of SSR and AIC are obtained for different assumptions on $\beta_i$. This result suggests that the two forms of $\beta_i$ used here do not affect the goodness of fits when the two viral infection modes are both considered. The reason is that in this case, cell-to-cell transmission predominates the virus infection (see relative contributions in next section). Thus, the infection rate by cell-free virus plays a very minor role in the fitting.

In conclusion, including the two virus infection modes (i.e. sequential cell-free virus infection and cell-to-cell transmission) is shown to be able to generate bimodal distribution of the proviral copy number observed in experiment (Jung et al., 2002). It follows from Table 1 that model (1) with $f_i = \binom{8}{i}\theta^i(1-\theta)^{8-i}$ generates smaller values of SSR and AIC for each patient than the model with $f_i = \frac{e^{-\gamma}}{1-e^{-\gamma}}\frac{\gamma^i}{i!}$. Therefore, the model with binomial distribution of $f_i$ provides a better fit to the data than the model with the modified Poisson distribution of $f_i$. Parameter values corresponding to the best-fit curves are given in Table 2.

### 3.2. Contributions to infection from different routes

Some previous studies have suggested that cell-to-cell transmission of HIV is more efficient than cell-free virus infection (Johnson and Huber, 2002; Mazurov et al., 2010; Dimitrov et al., 1993; Chen et al., 2007; Iwami et al., 2015; Graw and Perelson, 2016). Using data from tissue culture experiments, Dimitrov et al. (1993) found that the infectivity of cell-to-cell transmission is 100 to 1000 times greater than that of cell-free virus infection. A significant difference in the infectivity was also observed in an ultrasensitive, fluorescent virus transfer assay (Chen et al., 2007). In another study, it was shown that cell-associated virus transmission is 18,000-fold more efficient than cell-free virus infection. Using coupled experimental and mathematical investigation, Iwami et al. (2015) concluded that cell-to-cell infection accounts for approximately 60% of total infection. In our model, the multiplicity of infection per cell is caused by cell-to-cell transmission and sequential cell-free virus infection. Whether the multiple

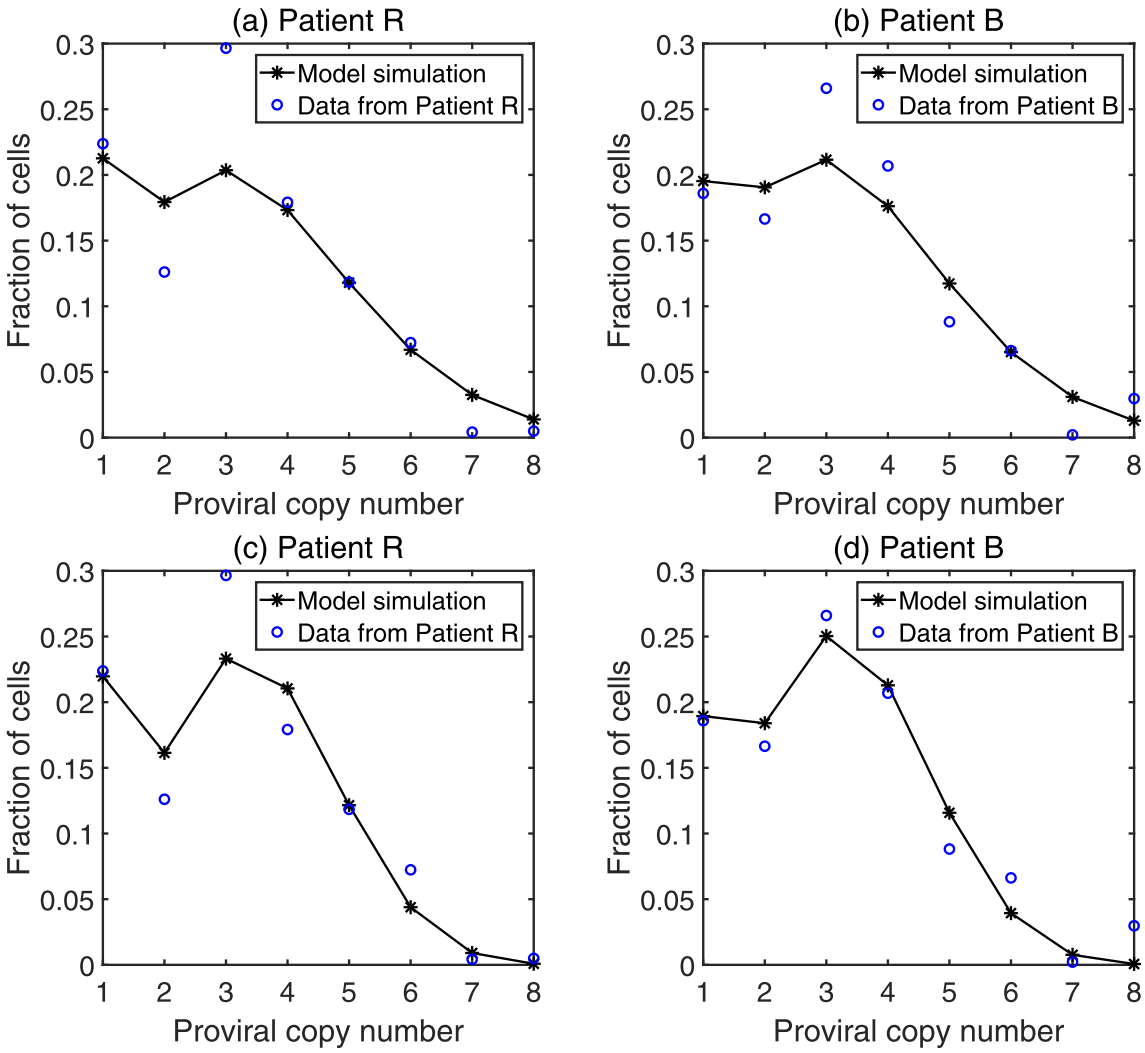


**Fig. 5.** Fitting using the model with both sequential virus infection and cell-to-cell transmission. The infection rate is assumed to be a decreasing function of the MOI, i.e. $\beta_i = \beta_1 \times 0.7^{i-1}$. Other descriptions are the same as in Fig. 4.

**Table 2**
Best-fit parameter estimates.*

| Patient | $\beta$ $(ml/day)$ | $\kappa$ $(ml/day)$ | $\theta$ (no unit) | $p$ $(cell^{-1}day^{-1})$ |
|---|---|---|---|---|
| R | $1.69 \times 10^{-8}$ | $5.6 \times 10^{-6}$ | 0.42 | 1414 |
| B | $1.3 \times 10^{-8}$ | $5.31 \times 10^{-6}$ | 0.4 | 1186 |
| **Mean** | $1.495 \times 10^{-8}$ | $5.455 \times 10^{-6}$ | 0.41 | 1300 |
| **SD** | $2.758 \times 10^{-9}$ | $2.051 \times 10^{-7}$ | 0.014 | 161.22 |

* Best-fit estimates of the infection rate by cell-free virus ($\beta$), rate of cell-to-cell transmission ($\kappa$), probability that a viral genome is successfully transmitted via cell-to-cell transmission ($\theta$), and viral production rate (p) obtained by fitting the model prediction to the proviral distribution data from two patients (Jung et al., 2002).

infection affects the evaluation of relative contributions from the two virus spread modes remains unclear.

We evaluate the relative contributions from these two modes of viral spread, which are given by the following two ratios

$$\frac{\beta_1 T(t)V(t)}{\beta_1 T(t)V(t) + \kappa T(t)I(t)} \quad \textit{and} \quad \frac{\kappa T(t)I(t)}{\beta_1 T(t)V(t) + \kappa T(t)I(t)},$$

where $I(t) = \sum_{i=1}^{n} I_i(t)$ represents the total number of infected cells at time t. Using model (1) with the means of best-fit parameter estimates listed in Table 2, we plot the relative contributions to CD4+ T cell infection in Fig. 6. At the initial stage of infection, there is only cell-free virus infection and no cell-to-cell transmission. This is because the host is initially infected only with free viruses. During the first 10 days of infection, the relative contribution from cell-free virus infection decreases, whereas the contribution from cell-to-cell transmission increases (see Fig. 6). As the infection progresses, model (1) eventually converges to the infected steady state $E^*$ (because the basic reproduction number $\mathcal{R}_0$ is greater than one with the best-fit parameter values), in which the concentrations of virus and cells remain unchanged. Thus, two curves representing the relative contributions to T cell infection also converge to constant values. The cell-to-cell transmission contributes about 87%

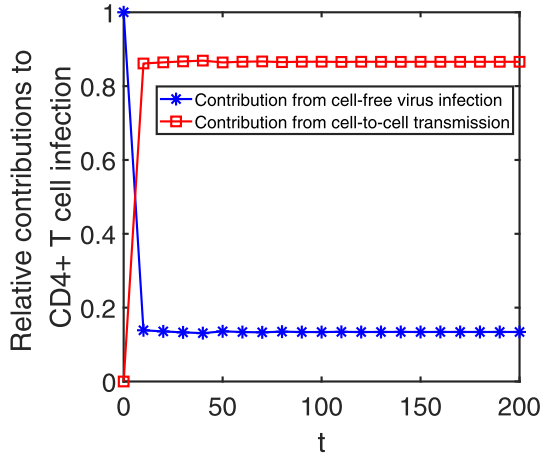


**Fig. 6.** The relative contributions to CD4+ T cell infection from cell-free virus infection and cell-to-cell transmission. Infection mainly comes from cell-free virus infection in the early stage of infection (before day 5). As the infection continues, the contribution from cell-associated transmission increases and converges to 87%. The values of parameters $\beta$, $\kappa$, $\theta$ and $p$ are the means of best-fit estimates listed in Table 2. Other parameter values are the same as in Fig. 2.

to the total infection. In other words, the contribution to CD4+ T cell infection from cell-to-cell transmission is approximately 6.5-fold higher than that from cell-free virus infection. This is consistent with previous estimate by Dixit and Perelson (2004) (i.e. approximately 90% of infection is caused by cell-to-cell transmission).

### 3.3. Relative contributions from infected cells with single and multiple proviruses

To provide further information on the dynamics of HIV multiple infections, we divide infected cells into two classes (i.e. infected cells with single provirus $I_1(t)$ and with multiple proviruses $I(t)-I_1(t)$). We study the relative contributions from these two classes of infected cells to CD4+ T cell infection. The two ratios are given by $I_1(t)/I(t)$ and $1-I_1(t)/I(t)$. As described in Fig. 7(a), only infected cells with single provirus exist at time 0. Thereafter, the contribution from infected cells with single provirus rapidly declines to 19% and the contribution from infected cells with multiple proviruses increases to 81%. Between days 1 and 5, the proportions of the two classes of infected cells in total infected cells stabilize at constant values, which means that the two classes of cells change with the same rate or remain unchanged. The same prediction also occurs after 20 days of infection. The contributions from infected cells with single and multiple proviruses stay at 21% and 79%, respectively. These values are in good agreement with *in vivo experimental results* (Jung et al., 2002).

We find in Fig. 7(a) that the relative contributions exhibit two dominant phases between day 5 and day 20. For the contribution from the infected cells with multiple proviruses, there is an initial increase, followed by a decline. An opposite change is observed in the contribution from the infected cells with single provirus (Fig. 7a). To further understand these phases, we plot the dynamics of the infected cells with single and multiple proviruses from day 5 to day 20 in Fig. 7(b, c). The simulations show that two types of infected cells both increase initially, followed by a decrease. It still cannot explain the two-phase change of the relative contributions shown in Fig. 7(a). To address this problem, we plot the ratio of infected cells with single provirus to infected cells with multiple proviruses (i.e. $\frac{I_1(t)}{I(t)-I_1(t)}$), shown in Fig. 7(d). This ratio declines slowly in the first phase and reaches the nadir at day 8.4, followed by an increase. Combining the results obtained in Fig. 7(b)–(d), we can describe the detailed dynamics of two types of infected cells between day 5 and day 20. When $t \in [5, 5.8]$, both infected cells with single and multiple proviruses increase, but the former has a slower increase rate. Compared with the infected cells with multiple proviruses, infected cells with single provirus decline at a faster rate from day 5.8 to day 8.4, and decline at a slower rate from day 8.4 to day 20. This explains the two-phase change in the relative contributions of two classes of infected cells, as shown in Fig. 7 (a).

In the zoom-in figure of Fig. 7(a), infected cells with multiple proviruses increase exponentially from time 0 to about 22% of total infected cells two hours after infection. Such a rapid increase may not appear in humans or animals after viral exposure. In our model, we did not include factors such as the eclipse phase and viral relocation from tissues to blood, which affect virus dynamics in blood in the early stage of infection *in vivo*. However, neglecting these factors allows us to compare the result with that obtained by Ito et al. Ito et al., 2018. Target cells were exposed to the virus inoculum for 2 h in an *in vitro* experiment Ito et al., 2018. By numerical simulation with best-fitted parameter values, they found that the multiply infected cells rapidly occurred and accumulated in up to 10% of all target cells within the time. Our prediction is more than twice the value in Ito et al., 2018. The difference may be due to the inclusion of cell-to-cell transmission in our model. It contributes 6.5 times higher to T cell infection than cell-free virus infection (Fig. 6).

To determine the role of cell-to-cell transmission in HIV multiple infections, we study the relative contributions from the two infection modes to two types of infected cells (i.e. infected cells with single and multiple proviruses) in Fig. 7(e). These relative contributions can be obtained by multiplying the contributions from two types of infected cells by their contributions from two infection routes. For example, the proportion of cells with single provirus infected by free virus in the total infected cells (blue dotted line in Fig. 7(e)) can be calculated by $\frac{I_1(t)}{I(t)} \times \frac{\beta_1 T(t) V(t)}{\beta_1 T(t) V(t) + \kappa T(t) I(t)}$. Fig. 7 (e) shows that the cells with single provirus infected by free virus are the most abundant population at the initial stage of infection. After that, the contributions from cell-to-cell transmission and multiple infections to the total infected cells increase rapidly. All curves oscillate and finally converge to the respective steady states. Infected cells with multiple proviruses caused by cell-to-cell transmission become the largest contributor to the total infection, accounting for about 70% of total infected cells. The infected cells with single provirus from cell-to-cell transmission account for 18% of T cell infection. The remaining of T cell infection comes from cell-free virus infection (the infected cells with single and multiple proviruses are 2% and 10%, respectively). These results suggest that cell-to-cell transmission has a larger impact on multiple infection of cells than cell-free virus infection.

## 4. Discussion

HIV multiple infection of cells has been observed both *in vitro* and *in vivo* (Dang et al., 2004; Portillo et al., 2011; Dixit and Perelson, 2005; Ito et al., 2017; Remion et al., 2016; Graw and Perelson, 2016), which greatly facilitates viral recombination (Bretscher et al., 2004; Kouyos et al., 2009; Law et al., 2016; Cromer et al., 2016). The biological processes underlying multiple infections are still not fully understood. Dixit and Perelson (2004) suggested that multiple infections could be explained by cell-to-cell transmission, in which multiple viral genomes are transmitted simultaneously in an infectious contact of an infected cell with a target cell. They also proposed another mechanism, that is, sequential cell-free virus infection to explain multiple infection (Dixit and Perelson, 2005). In this paper, we developed a mathematical model

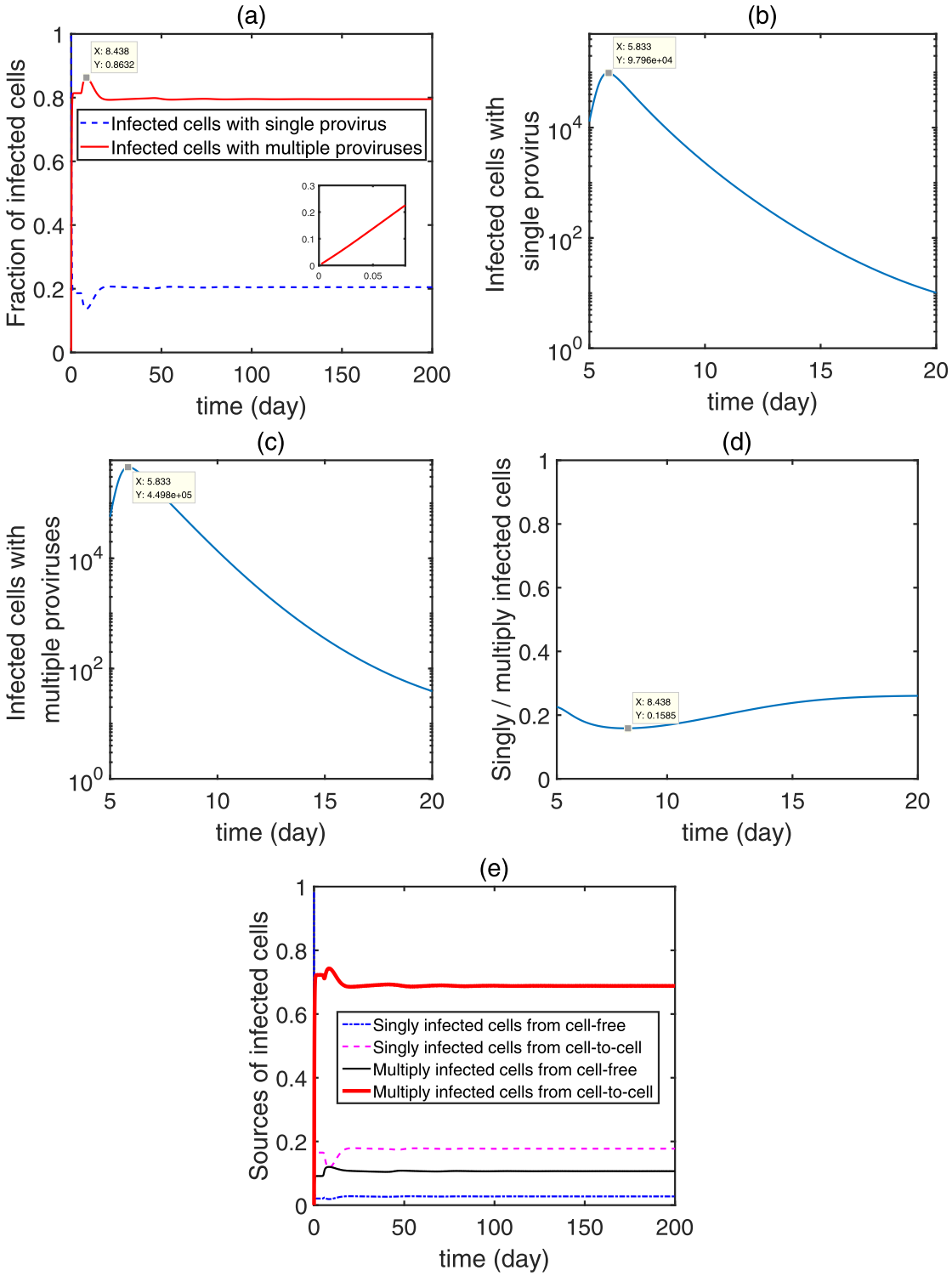


**Fig. 7.** Dynamics of infected cells with single and multiple proviruses during cell-to-cell transmission and cell-free virus infection. (a) The relative contributions to total infected cells from cells with single and multiple proviruses; (b) The dynamics of infected cells with single provirus; (c) The dynamics of infected cells with multiple proviruses; (d) The ratio of singly infected cells to multiply infected cells; (e) The relative contributions to total infection cells from two classes of infected cells (i.e. infected cells with single and multiple proviruses) and two modes of viral infection. The parameter values are the same as those in Fig. 6.

that includes both cell-to-cell transmission and sequential virus infection to study HIV multiple infection. We found that cell-to-cell transmission or sequential cell-free virus infection alone does not capture the bimodal distribution of the proviral copy number in experiment (Jung et al., 2002), as shown in Figs. 2 and 3. By fitting the full model to the experimental data, we show that both cell-to-cell transmission and sequential virus infection are needed to explain the distribution data of multiple infections (Figs. 4 and 5). These results further confirm the rationality of the model (1).

In the previous studies (Ito et al., 2018; Dixit and Perelson, 2005), the researchers showed that the frequencies of the infected cells with single provirus are higher than that of infected cells with

multiple proviruses. However, only cell-free virus infection was considered in these studies. In our model, we showed that both cell-free virus infection and cell-to-cell transmission are needed to explain the bimodal distribution of proviral data. The result agrees with the experimental result *in vivo* that more than 75% of HIV-infected cells in spleen harbor two or more proviruses (Jung et al., 2002). We also found that infected cells with multiple proviruses via cell-to-cell transmission account for approximately 70% of T cell infection (Fig. 7e). This further suggests that cell-to-cell transmission is an indispensable factor of studying multiple infections of cells.

Cell-to-cell transmission might be more effective in establishing the infection of cells compared with cell-free virus infection. As shown in Case C of Section 3.1, the values of SSR and AIC are not affected by the different assumptions on the sequential cell-free virus infection rate $\beta_i$ (either constant or decreasing) when the two viral transmission modes are both considered. We also simulated an extreme case in which infected CD4+ T cells are completely down-regulated and cannot be sequentially infected by cell-free virus (i.e. $\beta_1 > 0$ and $\beta_i = 0$ for $i = 2, 3, \ldots, 8$). We obtained very similar results. If there is only cell-free virus infection, then there are only infected cells with MOI = 1 (similar to the fitting in Fig. 2), which cannot explain the bimodal distribution data. When both transmission routes are included, then the fitting, the SSR and AIC values are very similar to Fig. 5. This is not surprising because in this case cell-to-cell transmission predominates the virus infection (see Fig. 6) and the fitting is not sensitive to the cell-free virus infection rate.

Although we cannot characterize the cell-free virus infection rates $\beta_i$ $(i = 1, 2, \ldots, 8)$ from available data, sequential cell-free virus infection may still be important in explaining viral recombination. A large number of recombinants were found in the splenocytes of the two patients in Jung et al. (2002). A prerequisite for recombination is the formation of heterozygotes, i.e. two copackaged RNA molecules with different genes (Dang et al., 2004). The occurrence of heterozygous virions requires the infection of a single cell by at least two different viral strains. Sequential cell-free virus infection may play a more important role than cell-to-cell transmission in generating genetic diversity. During sequential cell-free virus infection, viruses that have different genotypic and phenotypic properties are more likely to be brought into the same target cell. In contrast, multiple genomes acquired from a single cell-associated contact are expected to be of the same genotype. The relationship between the genetic variability in single target cell and two viral transmission modes has been quantitatively studied by Dixit and Perelson (2004). They estimated that the probabilities of target cell acquiring different genomes are $\sim$0.7 for sequential infection and $\sim$0.01 for cell-to-cell transmission. After sequential cell-free virus infection builds up a set of different genotypes in the same target cell, both sequential cell-free virus infection and cell-to-cell transmission would be consistent in establishing the diversity of proviruses.

In the model, the viral production rate $p$, the death rate of infected cells $\delta$, and the infection rate via cell-to-cell transmission $\kappa$ were assumed to be independent of the host cell's MOI, as used in some existing studies (Komarova et al., 2013; Dixit and Perelson, 2005; Althaus and Boer, 2012). They may depend on the cell's MOI. For example, different viral production rates were used for infected cells with different MOIs (Asatryan et al., 2015). We assumed a constant viral production rate for all infected cells due to the following considerations. First, virus cannot replicate without the machinery and metabolism of the host cell. Thus, viral production is more likely to be limited by cellular rather than viral factors. It is less likely that the large number of new viruses produced by an infected cell depends on the cell's MOI. Second, there are only limited data on the frequencies of cells with different MOI. Assuming different viral production rates will make the calculation of the frequencies at the steady state and the fitting to data challenging. Lastly, Dixit and Perelson (2005) found that the basic model without considering multiple infections, such as the models in Perelson (2002) and Nowak and May (2000), is able to successfully describe viral load dynamics in HIV infected individuals (Perelson et al., 1996, 1997; Wei et al., 1995). This is because virus production is assumed to depend on the number of infected cells rather than the number of integrated proviruses. When the viral production rate is independent of MOI, the cells (uninfected and total infected cells) and virus dynamics are decoupled from the dynamics of multiple infections. Thus, the evolutions of cells and virus are identical to that predicted by the basic model. Replacing multiply infected cells with singly infected cells do not alter viral dynamics. It is unclear whether assuming the death rate $\delta$ and cell-to-cell transmission rate $\kappa$ depend on the MOI generates richer dynamics, such as the bistability shown in Komarova and Wodar (2013).

In summary, our model incorporates the key forces that govern HIV multiple infections and quantitatively captures the experimental observations of proviral genomes in HIV-infected splenocytes. By numerical simulations with best-fitted parameter values, we estimate the relative contributions of two viral transmission modes to T cell single and multiple infections, which would otherwise be difficult to be determined by experiments. Both cell-to-cell transmission and cell-free virus infection are driving forces of multiple viral infection. If these virions are genetically distinct, then the emergence of viral recombination is inevitable. Some recombinants bring different advantageous alleles into a single genome, and thus may escape from host immunity and multidrug antiviral therapy. As a consequence, recombination speeds up the rate of HIV evolution. This study provides a modeling framework for revisiting some critical issues such as the recombination, emergence of drug resistance and viral evolution during HIV infection and treatment.

## CRediT authorship contribution statement

**Ting Guo:** Conceptualization, Methodology, Formal analysis, Writing - original draft. **Zhipeng Qiu:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Supervision. **Kosaku Kitagawa:** Methodology, Writing - review & editing. **Shingo Iwami:** Methodology, Writing - review & editing. **Libin Rong:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was finished when the first two authors visited the Department of Mathematics at University of Florida in 2018-2019. T. Guo was supported by the NSFC grant (12071217, 11971232), the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX20_0243), and the CSC (201806840119). Z. Qiu was supported by the NSFC grants (12071217, 11671206). L. Rong was supported by the NSF grants DMS-1758290 and DMS-1950254.

## Appendix A

**Proof of Theorem 2.2.** We first prove the global stability of the infection-free equilibrium $E_0$ when $\mathcal{R}_0 \leqslant 1$. Define a Lyapunov function

$$L = \left(T - \frac{\Lambda}{\mu} - \frac{\Lambda}{\mu}\ln\frac{T}{\frac{\Lambda}{\mu}}\right) + \mathcal{R}_0\sum_{j=1}^{n} I_j + \frac{\beta_1}{c}\frac{\Lambda}{\mu}V.$$

Differentiating $L$ along the solutions of system (1) yields

$$\begin{aligned}\frac{dL}{dt} &= \left(1 - \frac{\Lambda}{\mu T}\right)\left(\Lambda - \mu T - \beta_1 TV - \kappa T\sum_{j=1}^{n} I_j\right)\\ &\quad + \mathcal{R}_0\left(\beta_1 TV + \kappa T\sum_{j=1}^{n} I_j - \delta\sum_{j=1}^{n} I_j\right) + \frac{\beta_1}{c}\frac{\Lambda}{\mu}\left(p\sum_{j=1}^{n} I_j - cV\right). \end{aligned} \tag{11}$$

After collecting and rearranging terms, we obtain that

$$\frac{dL}{dt} = -\frac{\mu}{T}\left(\frac{\Lambda}{\mu} - T\right)^2 + (\mathcal{R}_0 - 1)\left(\beta_1 TV + \kappa T\sum_{j=1}^{n} I_j\right). \tag{12}$$

Since $\mathcal{R}_0 \leqslant 1$, we obtain that $\frac{dL}{dt} \leqslant 0$. Therefore, all limit points are contained in the largest invariant subset of

$$G = \left\{(T, I_1, \ldots, I_n, V) : \frac{dL}{dt} = 0\right\}.$$

It is clear that $\frac{dL}{dt} = 0$ if and only if $T = \frac{\Lambda}{\mu}, I_i = 0$ $(i = 1, 2, \ldots, n)$, and $V = 0$. Thus, the only compact invariant subset of $G$ is the singleton set $\{E_0\}$. By LaSalle's Invariance Principle, we know that $E_0$ is globally asymptotically stable when $\mathcal{R}_0 \leqslant 1$.

Next, we show the global stability of the positive equilibrium $E^*$. The system (1) can be reduced into

$$\begin{cases} \frac{dT}{dt} = \Lambda - \mu T(t) - \beta_1 T(t)V(t) - \kappa T(t)I(t),\\ \frac{dI}{dt} = \beta_1 T(t)V(t) + \kappa T(t)I(t) - \delta I(t),\\ \frac{dV}{dt} = pI(t) - cV(t). \end{cases} \tag{13}$$

Clearly, the system (13) has a unique positive equilibrium $E^*(T^*, I^*, V^*)$ when $\mathcal{R}_0 > 1$, where $I^* = \sum_{i=1}^{n} I_i^*$. Define the following Lyapunov function

$$L^* = (T - T^* - T^*\ln\frac{T}{T^*}) + (I - I^* - \ln\frac{I}{I^*}) + \frac{\beta_1 T^*}{c}(V - V^* - V^*\ln\frac{V}{V^*}). \tag{14}$$

The derivative of $L^*$ along the solutions of (13) is

$$\begin{aligned}\frac{dL^*}{dt} &= \left(1 - \frac{T^*}{T}\right)(\Lambda - \mu T - \beta_1 TV - \kappa TI)\\ &\quad + \left(1 - \frac{I^*}{I}\right)(\beta_1 TV + \kappa TI - \delta I) + \left(1 - \frac{V^*}{V}\right)(pI - cV)\\ &= -\mu\frac{(T - T^*)^2}{T} + \beta_1 T^*V^*\left(3 - \frac{T^*}{T} - \frac{TI^*V}{T^*V^*I} - \frac{IV^*}{VI^*}\right)\\ &\quad + \kappa T^*I^*\left(2 - \frac{T^*}{T} - \frac{T}{T^*}\right), \end{aligned} \tag{15}$$

where we used the steady state equalities

$$\Lambda = \mu T^* + \beta_1 T^*V^* + \kappa T^*I^*, \qquad \delta I^* = \beta_1 T^*V^* + \kappa T^*I^*, \qquad pI^* = cV^*.$$

We obtain that $\frac{dL^*}{dt} \leq 0$. Furthermore, $\frac{dL^*}{dt} = 0$ if and only if $T = T^*$, $I = I^*$ and $V = V^*$. This implies that $T(t) \to T^*, I(t) \to I^*$ and $V(t) \to V^*$ as $t \to +\infty$. Thus, the limit system of (1) is

$$\begin{cases} \frac{dI_1(t)}{dt} = \beta_1 T^*V^* - \beta_2 V^*I_1(t) + f_1\kappa T^*I^* - \delta I_1(t),\\ \frac{dI_i(t)}{dt} = \beta_i V^*I_{i-1}(t) - \beta_{i+1}V^*I_i(t) + f_i\kappa T^*I^* - \delta I_i(t),\\ i = 2, 3 \ldots, n-1,\\ \frac{dI_n(t)}{dt} = \beta_n V^*I_{n-1}(t) + f_n\kappa T^*I^* - \delta I_n(t). \end{cases} \tag{16}$$

When $\mathcal{R}_0 > 1$, the linear system (16) has only one positive equilibrium $\tilde{E}^* = (I_1^*, I_2^*, \ldots, I_n^*)$, where $I_1^*, I_2^*, \ldots, I_n^*$ are given in (9). Calculating the Jacobian matrix of (16) at $\tilde{E}^*$, we have

$$H(\tilde{E}^*) = \begin{pmatrix} -\beta_2 V^* - \delta & 0 & 0 & \cdots & 0 & 0\\ \beta_2 V^* & -\beta_3 - \delta & 0 & \cdots & 0 & 0\\ 0 & \beta_3 V^* & -\beta_4 - \delta & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & \beta_n V^* & -\delta \end{pmatrix}_{n\times n}. \tag{17}$$

Since (17) is a lower triangular matrix and its diagonal entries are negative, all eigenvalues of matrix (17) have negative real parts. Thus, $\tilde{E}^*$ is locally asymptotically stable. Using Theorem 1.2.2 of Zhao (2017), it follows that the positive equilibrium $E^*$ is globally asymptotically stable when $\mathcal{R}_0 > 1$. This completes the proof of Theorem 2.2.

## References


Agosto, L.M., Uchil, P., Mothes, W., 2015. HIV cell-to-cell transmission: effects on pathogenesis and antiretroviral therapy. Trends Microbiol. 23, 289–295.

Althaus, C.L., Boer, R.J.D., 2012. Impaired immune evasion in HIV through intracellular delays and multiple infection of cells. Proc. R. Soc. B 279, 3003–3010.

Asatryan, A., Wodarz, D., Komarova, N.L., 2015. New virus dynamics in the presence of multiple infection. J. Theor. Biol. 377, 98–109.

Bofill, M., Janossy, G., Lee, C.A., et al., 1992. Laboratory control values for CD4 and CD8 T lymphocytes. implications for HIV-1 diagnosis. Clin. Exp. Immunol. 88, 243–252.

Bretscher, M., Althaus, C., Müller, V., et al., 2004. Recombination in HIV and the evolution of drug resistance: for better or for worse?. Bioessays 26, 180–188.

Chen, J., Dang, Q., Unutmaz, D., et al., 2005. Mechanisms of nonrandom human immunodeficiency virus type 1 infection and double infection: preference in virus entry is important but is not the sole factor. J. Virol. 79, 4140–4149.

Chen, P., Hübner, W., Spinelli, M., et al., 2007. Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained env-dependent neutralization-resistant virological synapses. J. Virol. 81, 12582–12595.

Cromer, D., Grimm, A.J., Schlub, T.E., et al., 2016. Estimating the in-vivo HIV template switching and recombination rate. AIDS 30, 185–192.

Dang, Q., Chen, J.B., Unutmaz, D., et al., 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc. Natl. Acad. Sci. USA 101, 632–637.

den Driessche, P.V., Watmough, J., 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 180, 29–48.

Dimitrov, D.S., Willey, R.L., Sato, H., et al., 1993. Quantitation of human immunodeficiency virus type 1 infection kinetics. J. Virol. 67, 2182–2190.

Dixit, N.M., Perelson, A.S., 2004. Multiplicity of human immunodeficiency virus infections in lymphoid tissue. J. Virol. 78, 8942–8945.

Dixit, N.M., Perelson, A.S., 2005. HIV dynamics with multiple infections of target cells. Proc. Natl. Acad. Sci. USA 102, 8198–8203.

Graw, F., Perelson, A.S., 2016. Modeling viral spread. Annu. Rev. Virol. 3, 555–572.

Guo, T., Qiu, Z., 2019. 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, 6822–6841.

Hill, A.L., Rosenbloom, D.S., Nowak, M.A., et al., 2018. Insight into treatment of HIV infection from viral dynamics models. Immunol. Rev. 285, 9–25.

Hübner, W., McNerney, G., Chen, P., et al., 2009. Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science 323, 1743–1747.

Ito, Y., Remion, A., Tauzin, A., et al., 2017. Number of infection events per cell during HIV-1 cell-free infection. Sci. Rep. 7, 6559.

Ito, Y., Tauzin, A., Remion, A., et al., 2018. Dynamics of HIV-1 coinfection in different susceptible target cell populations during cell-free infection. J. Theor. Biol. 455, 39–46.

Iwami, S., Takeuchi, J.S., Nakaoka, S., et al., 2015. Cell-to-cell infection by HIV contributes over half of virus infection. eLife 4. e08150.

Johnson, D.C., Huber, M.T., 2002. Directed egress of animal viruses promotes cell-to-cell spread. J. Virol. 76, 1–8.
Jolly, C., Sattentau, Q.J., 2004. Retroviral spread by induction of virological synapses. Traffic 5, 643–650.
Jung, A., Maier, R., Vartanian, J.P., et al., 2002. Recombination: multiply infected spleen cells in HIV patients. Nature 418, 144.
Komarova, N.L., Wodar, D., 2013. Virus dynamics in the presence of synaptic transmission. Math. Biosci. 242, 161–171.
Komarova, N., Levy, D., Wodarz, D., 2013. Synaptic transmission and the susceptibility of HIV infection to anti-viral drugs. Sci. Rep. 3, 2103.
Kouyos, R., Fouchet, D., Bonhoeffer, S., 2009. Recombination and drug resistance in HIV: population dynamics and stochasticity. Epidemics 1, 58–69.
Kumberger, P., Frey, F., Schwarz, U.S., et al., 2016. Multiscale modeling of virus replication and spread. FEBS Lett. 590, 1972–1986.
Kumberger, P., Durso-Cain, K., Uprichard, S.L., et al., 2018. Accounting for space-quantification of cell-to-cell transmission kinetics using virus dynamics models. Viruses 10. e200.
Lai, X., Zou, X., 2014. Modelling HIV-1 virus dynamics with both virus-to-cell infection and cell-to-cell transmission. SIAM J. Appl. Math. 74, 898–917.
Law, K.M., Komarova, N.L., Yewdall, A.W., et al., 2016. In vivo HIV-1 cell-to-cell transmission promotes multicopy micro-compartmentalized infection. Cell Rep. 15, 2771–2783.
Levy, D.N., Aldrovandi, G.M., Kutsch, O., et al., 2004. Dynamics of HIV-1 recombination in its natural target cells. Proc. Natl. Acad. Sci. USA 101, 4204–4209.
Mazurov, D., Ilinskaya, A., Heidecker, G., et al., 2010. Quantitative comparison of HTLV-1 and HIV-1 cell-to-cell infection with new replication dependent vectors. PLoS Path. 6. e1000788.
Mohri, H., Bonhoeffer, S., Monard, S., et al., 1998. Rapid turnover of T lymphocytes in SIV-infected rhesus macaques. Science 279, 1223–1227.
Mothes, W., Sherer, N.M., Jin, J., et al., 2010. Virus cell-to-cell transmission. J. Virol. 84, 8360–8368.
Nowak, M.A., May, R.M., 2000. Virus Dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford.
Perelson, A.S., 2002. Modelling viral and immune system dynamics. Nat. Rev. Immunol. 2, 28–36.
Perelson, A.S., Nelson, P.W., 1999. Mathematical analysis of HIV-1 dynamics in vivo. SIAM Rev. 41, 3–44.
Perelson, A.S., Ribeiro, R.M., 2013. Modeling the within-host dynamics of HIV infection. BMC Biol. 11, 96.
Perelson, A.S., Neumann, A.U., Markowitz, M., et al., 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271, 1582–1586.
Perelson, A.S., Essunger, P., Cao, Y., et al., 1997. Decay characteristics of HIV-1-infected compartments during combination therapy. Nature 387, 188–191.
Portillo, A.D., Tripodi, J., Najfeld, V., et al., 2011. Multiploid inheritance of HIV-1 during cell-to-cell infection. J. Virol. 85, 7169–7176.
Remion, A., Delord, M., Hance, A.J., et al., 2016. Kinetics of the establishment of HIV-1 viral interference and comprehensive analysis of the contribution of viral genes. Virol 487, 59–67.
Rong, L., Perelson, A., 2009. Modeling HIV persistence, the latent reservoir, and viral blips. J. Theor. Biol. 260, 308–331.
Rong, L., Perelson, A.S., 2009. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. PLoS Comput. Biol. 5. e1000533.
Sigal, A., Kim, J.T., Balazs, A.B., et al., 2011. Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. Nature 477, 95–98.
Stafford, M.A., Corey, L., Cao, Y., et al., 2000. Modeling plasma virus concentration during primary HIV infection. J. Theor. Biol. 203, 285–301.
Suryavanshi, G.W., Dixit, N.M., 2007. Emergence of recombinant forms of HIV: dynamics and scaling. PLoS Comput. Biol. 3. e205.
Wang, X., Rong, L., 2019. HIV low viral load persistence under treatment: Insights from a model of cell-to-cell viral transmission. Appl. Math. Lett. 94, 44–51.
Wang, X., Tang, S., Song, X., et al., 2017. Mathematical analysis of an HIV latent infection model including both virus-to-cell infection and cell-to-cell transmission. J. Biol. Dyn. 11 (sup2), 455–483.
Wei, X., Ghosh, S.K., Taylor, M.E., et al., 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature 373, 117–122.
Zhao, X., 2017. Dynamical Systems in Population Biology. Springer-Verlag, New York.
Zitzmann, C., Kaderali, L., 2018. Mathematical analysis of viral replication dynamics and antiviral treatment strategies: from basic models to age-based multiscale modeling. Front. Microbiol. 9, 1546.