Contribution of HIV-1 genomes that do not integrate to the basic reproductive ratio of the virus

John Wei Lau[1,*], **David N. Levy**[2], and **Dominik Wodarz**[1,*]

[1]Department of Ecology and Evolutionary Biology, 321 Steinhaus Hall, University of California, Irvine

[2]Department of Basic Science, New York University College of Dentistry, 921 Schwartz Building, 345 East 24th Street, New York, NY 10010-9403

Abstract

Recent experimental data indicate that HIV-1 DNA that fails to integrate (from now on called uDNA) can by itself successfully produce infectious offspring virions in resting T cells that become activated after infection. This scenario is likely important at the initial stages of the infection. We use mathematical models to calculate the relative contribution of unintegrated and integrated viral DNA to the basic reproductive ratio of the virus, R_0, and the models are parameterized with preliminary data. This is done in the context of both free virus spread and transmission of the virus through virological synapses. For free virus transmission, we find that under preliminary parameter estimates, uDNA might contribute about 20% to the total R_0. This requires that a single copy of uDNA can successfully replicate. If the presence of more than one uDNA copy is required for replication, uDNA does not contribute to R_0. For synaptic transmission, uDNA can contribute to R_0 regardless of the number of uDNA copies required for replication. The larger the number of viruses that are successfully transmitted per synapse, however, the lower the contribution of uDNA to R_0 because this increases the chances that at least one virus integrates. Using available parameter values, uDNA can maximally contribute 20% to R_0 in this case. We argue that the contribution of uDNA to virus reproduction might also be important for continued low level replication of HIV-1 in the presence of integrase inhibitor therapy. Assuming a 20% contribution of uDNA to the overall R_0, our calculations suggest that $R_0=1.6$ in the absence of virus integration. While these are rough estimates based on preliminary data that are currently available, this analysis provides a framework for future experimental work which should directly measure key parameters.

*Corresponding authors: John Lau: wlau1@uci.edu, Tel: 415-425-5426, Dominik Wodarz: dwodarz@uci.edu, Tel: 949-824-2531.

Introduction

Human immunodeficiency virus-1 (HIV-1) infects cells of the immune system, mainly CD4+ T cells and antigen presenting cells. The virus has an RNA genome, which upon infection is copied into DNA through reverse transcription. The DNA genome of the virus can subsequently integrate into the host genome, and this forms the template for transcription, eventually leading to the formation of new virus particles that are released from the infected cell. This cycle of infection and replication allows the virus to spread from cell to cell, leading to the exponential growth of the HIV-1 population during the acute phase of the infection. After initiation of immune responses and also depletion of many target cells during the acute phase (Zhang et al., 2004), the virus population settles around a steady state level during the chronic phase of infection, termed the set point viral load. The set point viral load is highly correlated with the speed at which infection progresses from the asymptomatic phase to AIDS (Lifson et al., 1997; Mellors et al., 1996).

Extensive clinical and experimental data document these dynamics both in vivo and during in vitro experiments. Mathematical models have been very useful to help interpret those data, to measure crucial parameters, and to generate testable hypotheses (Nowak and May, 2000; Perelson, 2002; Perelson and Ribeiro, 2013; Wodarz, 2006). One of the most fundamental measures is the basic reproductive ratio of the virus, defined as the average number of newly infected cells generated by a single infected cell (Anderson and May, 1991; Heffernan et al., 2005; Nowak and May, 2000). It has been estimated for HIV-1 in a variety of settings (Little et al., 1999; Nowak et al., 1997; Ribeiro et al., 2010), and has important implications for understanding disease progression as well as the response to anti-viral drug therapies (Bonhoeffer and Nowak, 1997; Bonhoeffer et al., 1997).

During the replication cycle of the virus, the step of integration is prone to failure, leading to the generation of unintegrated viral DNA genomes (uDNA). While some of the unintegrated DNA found in cells is simply pre-integration DNA that will eventually integrate, most viral DNA fails to integrate and remains extrachromosomal. Throughout the rest of this paper, we will use the term uDNA to refer to the latter, i.e. to viral DNA that fails to integrate. Pre-integration DNA is excluded from the current analysis and discussion, because it will eventually integrate. uDNA is present in both linear form, which is subject to exonuclease digestion at poorly defined rates, and circularized forms containing one and two long terminal repeats (1-LTR and 2-LTR circles). Circularized uDNA is highly stable and lost only through cell death or dilution during cell division. Gene expression has been detected from all forms of uDNA, linear and circular (Cara et al., 1996; Sloan and Wainberg, 2011; Trinite et al., 2013; Wu and Marsh, 2001).

According to the literature, uDNA is considered a replicative dead end and does not lead to the production of infectious offspring virus (e.g. (Butler et al., 2001)). Recent finding from two of us, however, indicate that the situation could be more complex (Gelderblom et al., 2008; Trinite et al., 2013). Thus, if a cell is coinfected with integrated and unintegrated virus, viral uDNA has been shown to successfully produce infectious offspring virus, indicating cooperative interactions (Gelderblom et al., 2008). Even more striking, our recent study has shown that cells that only contain viral uDNA (that fails to integrate) can

successfully produce infectious offspring virus (Trinite et al., 2013). This has been shown using an integration-defective mutant, as well as with wild-type virus in cells treated with an integrase inhibitor (Trinite et al., 2013). Virus production in uDNA-infected cells occurs at a rate that is small compared to integrated genomes (Trinite et al., 2013); however, cells infected only with uDNA also live significantly longer than productively infected cells that contain iDNA, thus increasing the total amount of virus produced by those cells during their life-span (manuscript in preparation). Successful uDNA replication, however, does not seem to occur under all conditions. It is observed specifically in resting cells that are infected and subsequently become activated (Trinite et al., 2013), which could be especially relevant to the initial stages of the infection, a time when many resting $CD4^+$ T cells are infected (Haase, 2011; Hladik and McElrath, 2008; Kinter et al., 2003; Li et al., 2005; Nishimura and Martin, 2011; Reilly et al., 2007; Zhang et al., 1999; Zhang et al., 2004) and the latent reservoir is established (Chun et al., 1998; Hatziioannou et al., 2014). Thus, uDNA could have importance for the ability of the virus to establish a persistent infection, reflected in the value of its basic reproductive ratio. Replication from uDNA might also play a role when virus load is reduced to very low levels during anti-viral therapies, and following treatment interruption, when latent viruses re-emerge.

This paper presents for the first time a mathematical model that allows successful virus replication both from cells containing integrated virus and from cells that only contain virus that fails to integrate. The aim of this study is to quantify the contribution of viral uDNA to the basic reproductive ratio of the virus, and thus to the ability of the virus to establish a persistent infection. We start with a standard virus dynamics model and include the possibility of uDNA replication. This model assumes that the virus spreads through the target cell population via the release of free virus particles. In this model, iDNA and uDNA replication contribute additively to the basic reproductive ratio of the virus, and their relative contributions are determined by the parameters describing the rate of virus production and cell death. This requires that a cell with a single copy of viral uDNA can produce offspring virus. While data indicate that this can indeed occur, it cannot currently be discounted that multiple uDNA copies are required to successfully produce offspring virus. In this case, uDNA replication does not contribute to the basic reproductive ratio of the virus, because multiple infection is a very unlikely event in the context of free virus transmission at relatively low virus loads. In addition to this model we also consider the assumption that the virus can spread directly from cell to cell, through the formation of virological synapses (Hubner et al., 2009). In this model, uDNA can contribute to the basic reproductive ratio of the virus even if multiple uDNA copies are required for successful replication. The reason is that synaptic transmission is well documented to deliver multiple infectious viruses, resulting in a high multiplicity of infection of cells (MOI) (Del Portillo et al., 2011). The exact contribution of uDNA to the basic reproductive ratio of the virus depends on the number of viruses that get transferred through the synapse. The larger this number, the higher the likelihood that a virus will integrate into the host genome, and thus the lower the contribution of uDNA to the basic reproductive ratio. In the context of available and preliminary parameters, the model suggests that uDNA replication might contribute 20% to the overall R_0 for free virus transmission, and maximally 20% for synaptic transmission, depending on the number of viruses that get successfully transmitted per synapse. While this

is a relatively small fraction, we estimate that the basic reproductive ratio of HIV-1 in the absence of integrated virus can be still greater than one in the context of these parameter values, potentially allowing the occurrence of persistent, low-level viral replication during treatment with integrase inhibitors. These insights are discussed in the context of data that examine virus dynamics during raltegravir intensification of HAART.

The basic mathematical model of virus dynamics with free virus transmission

To set the scene, we briefly summarize a basic model of virus dynamics that has been analyzed extensively in the literature. It assumes transmission of free virus released from cells. For a review, see (Nowak and May, 2000; Perelson, 2002; Perelson and Ribeiro, 2013; Wodarz, 2006). It considers the populations of susceptible, uninfected cells, x, infected cells, y, and free virus, v. It is given by the following set of ordinary differential equations, which are written in "dot form" (equivalent to dx/dt, etc).

$$\begin{aligned} \dot{x} &= \lambda - dx - \beta xv \\ \dot{y} &= \beta xv - ay \\ \dot{v} &= ky - uv \end{aligned} \quad (1)$$

Uninfected cells are generated with a rate λ, die with a rate d, and become infected by virus with a rate β. Infected cells die with a rate a, and produce free virus particles with a rate k. Free virus decays with a rate u. This model is characterized by two equilibria. If the virus fails to establish a persistent infection, the system converges to the following equilibrium: $x^{(0)} = \lambda/d; \ y^{(0)} = 0; \ v^{(0)} = 0$. If the virus population successfully establishes a persistent infection, the system converges to the following equilibrium.

$x^{(1)} = \dfrac{au}{\beta k}; y^{(1)} = \dfrac{\lambda\beta k - dau}{a\beta k}; v^{(1)} = \dfrac{\lambda\beta k - dau}{au\beta}$. Which outcome is observed depends on the basic reproductive ratio of the virus, $R_0 = \lambda\beta k/dau$. A persistent infection is established if $R_0 > 1$. For further details of this model, see (Nowak and May, 2000) and references therein.

Free virus transmission and uDNA Replication

Here, we consider an extension of the basic model that includes multiple infection of cells and takes into account the replication of both integrated and unintegrated viral genomes. Thus, we denote infected cells with i unintegrated and j integrated viruses by y_{ij}. Consistent with this notation, the variable y_{00} represents the population of uninfected target cells. The model is thus given by the following set of ordinary differential equations (also see schematic in Figure 1a).

$$\begin{aligned} \dot{y}_{0,0} &= \lambda - dy_{0,0} - \beta y_{0,0} v \\ \dot{y}_{i,j} &= (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j} \\ \dot{v} &= \sum_{i=0}^{N} \sum_{j=0}^{M} k_{ij} y_{ij} - uv \end{aligned} \quad (2)$$

As in the previous model (1), infection occurs with a rate β. Upon infection, integration of the viral genome occurs with a probability p, while uDNA capable of replication is generated with a probability 1−p. As explained above, we consider only viral DNA that will

not integrate (Trinite et al., 2013), rather than DNA that retains a functional pre-integration complex and might subsequently integrate. That is, pre-integration uDNA is not considered. The rest of the model is the same as the basic model discussed above. Kinetic parameters, such as the rate of virus production of infected cells and the rate of infected cell death can depend on the number of integrated and unintegrated viruses. Data indicate that cells infected with uDNA only replicate the virus with a slower rate but also die with a slower rate than cells infected with iDNA (manuscript in preparation). However, beyond this, there is no experimental evidence that the number of iDNA and uDNA copies in cells influences these parameters. Thus, we make the simplifying assumption that k_{ij} and a_{ij} can each have only two values. Cells with only unintegrated viruses produce virus at rate k_u and die at rate a_u. Cells with any amount of integrated viruses produce virus at a rate k_i and also die at a rate of a_i (whether uDNA is present or not). Note that in this model, the term iDNA refers to cells that are productively infected with integrated viral DNA. Latently infected cells are not taken into account here.

A couple of notes regarding the model structure: (i) The model includes a full multiple infection cascade, i.e. taking into account cells infected with i unintegrated and j integrated genomes. It could be argued that a simpler version of the model is sufficient, since only two types of infected cells are important to our analysis: those containing only uDNA and those containing iDNA. We have, however, chosen to explore the full model in order to avoid potential complications that can arise from simplified models (explored in Phan et al, submitted). (ii) uDNAs that fail to integrate can have different life-spans in the cells, depending on their exact form. The intracellular dynamics of uDNA (or iDNA) is not modeled here. We assume that the genomes considered in our models persist for the life-span of the infected cell.

As in the basic model of virus dynamics, this system is also characterized by two equilibria. In the absence of infection, the system converges to the trivial equilibrium, given by is $y_{00}=$ λ/d, $y_{ij}=0$, and $v=0$. If the virus establishes a persistent infection, the system converges to the internal equilibrium where all populations are greater than zero (not written down due to complexity of expressions).

Using standard methods (see Supplementary Material), we can calculate the threshold for a successful infection from Model (2), defined by the stability of the virus-free steady state. This gives rise to the basic reproductive ratio of the virus in model (2), which is given by

$$R_0 = p\frac{\lambda \beta k_i}{a_i u d} + (1-p)\frac{\lambda \beta k_u}{a_u u d}$$

This is similar to the expression for R_0 in the basic model (1), but has contributions from virus replication originating from iDNA (p) and uDNA ($1-p$). If we assume that integrated and unintegrated viruses contribute equally to virus replication and cell death, we can sum the terms and arrive at same expression for R_0 as in the basic model of virus dynamics, model (1). The same would be true if the burst size, defined as the average number of

viruses produced per infected cell, was the same for both unintegrated and integrated viruses.

The analysis so far has assumed that a single copy of uDNA alone can lead to successful virus replication. Let us now assume a more complex situation in which more than one uDNA copy needs to be present in a cell for virus replication to be possible in the absence of iDNA. While there is currently no evidence that this is a requirement, this scenario also cannot currently be discounted. In this case, the basic reproductive ratio of the virus becomes:

$$R_0 = p \frac{\lambda \beta k_i}{a_i u d}$$

In other words, unintegrated viruses do not contribute to establishing infection anymore. Intuitively, this can be explained as follows. At the beginning of infection, the number of uninfected cells is high while the number of viruses is relatively low. Thus, the chance that a cell is multiply infected is negligible. Since our current assumption requires multiple unintegrated viruses in a cell to start viral production, the cells infected with a single unintegrated virus are not yet productive at the very early stages of infection.

Synaptic virus transmission and uDNA Replication

Here, the model is modified to describe synaptic transmission of the virus. The model is given by the following set of ordinary differential equations (also see schematic, Figure 1b), which are based on previous work (Komarova and Wodarz, 2013; Komarova et al., 2012b; Komarova et al., 2013a; Komarova et al., 2013b).

$$\dot{y}_{0,0} = \lambda - d y_{0,0} - \beta y_{0,0} v - y_{0,0} \sum_{k=0}^{N} \sum_{l=0}^{M} y_{kl} \sum_{r=0}^{N} \sum_{s=0}^{M} \gamma_{rs}^{kl}$$

$$\dot{y}_{i,j} = (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j} + \sum_{k=0}^{N} \sum_{l=0}^{M} y_{kl} \left(\sum_{r=0}^{i} \sum_{s=0}^{j} y_{rs}^{kl} y_{i-r,j-s} - y_{ij} \sum_{r=0}^{N-i} \sum_{s=0}^{M-j} \gamma_{rs}^{kl} \right) \quad (3)$$

$$\dot{v} = \sum_{i=0}^{N} \sum_{j=0}^{M} k_{ij} y_{ij} - u v$$

for $i \ N, j \ M, i+j \ N$

Compared to model (2), it has one additional parameter: γ_{rs}^{kl}, which describes the probability that a cell with k unintegrated and l integrated viral genomes transmits r unintegrated and s integrated viruses (more precisely transmit genomes, *r* of which will be unintegrated and *s* will be integrated in the target cell). The γ term is an analog of the infectivity used in free virus transmission and is composed of 3 parts. Incorporated in this constant is the probability that a synapse is formed when two cells interact, the probability that r+s total viruses are successfully transmitted and the probability that r viral genomes are unintegrated while s viral genomes become integrated provirus in the target cell.

The equation describing the dynamics of uninfected cells, x, is identical to model (2) except with an added term for cell-to-cell transmission from other cells. The extra term describes

cells with k and l unintegrated and integrated viral genomes transferring r viruses that will be unintegrated genomes and s that will be integrated genomes in the target cell. The equation for infected cells is also identical to that used in model (2) with the addition of terms that describe cell-to- cell interactions. The first term for cell-to-cell interactions describes other cells gaining viruses to form cells of type y_{ij}. The second term for cell-to-cell interactions describes cells of type y_{ij} gaining viruses by cell-to-cell transmission.

In the above equations, we place a limit on the number of viruses allowed per cell to be N. We also further limit the number of integrated viruses to be M. Thus, we only model y_{ij} for indexes such that i+j ≤ M. For simplicity, we did not explicitly write the equations for the endpoints, i.e. y_{iM}, y_{Nj}, y_{NM}. These species cannot be further infected in these equations.

Note that consistent with the other models, the synaptic transmission model is also given by ordinary differential equations, which imply perfect mixing of cells. In tissues, source cells are likely to form synapses with target cells in their vicinity, which could add a spatial component to the dynamics. Imaging data in lymph nodes, however, indicate that a lot of cell movement and mixing occurs in tissue (Germain et al., 2012). Thus, while synapses might be formed between neighboring cells, ongoing mixing means that a cell has a high chance to interact with many other cells in the system. A large migration rate in spatial models has been shown to lead to mass-action properties (Komarova et al., 2012a; Thalhauser et al., 2010), and hence the ODE modeling approach is justified here.

In order to arrive at an analytic result for R_0 for the system with cell to cell transmission, we use some simplifying assumptions. We first explore the system with synaptic transmission only by simplifying and separating the infectivity term, $\gamma_{rs}{}^{kl}$, to its constituent parts.

We assume that the rates of cell-to-cell transmission are determined by the presence of iDNA in the source cell. If a cell contains one or more integrated viral genomes, the kinetics of virus production and cell death are dictated by the integrated virus, and the potential presence of uDNA in the same cell is not assumed to make a difference. On the other hand, in the absence of iDNA, virus production by infected cells and the death rate of infected cells are determined by the parameters of the uDNA infection. It is assumed that the rate of virus production depends only on the type, but not the number, of viruses in the infected cell. Under this assumption, the parameter $\gamma_{rs}{}^{kl}$ can be simplified to $\hat{\gamma}_{r,s}^u$ and $\hat{\gamma}_{r,s}^i$, for any k and l. Here, $\hat{\gamma}_{r,s}^u$ is the reduced transmission rate from cells with only unintegrated viral genomes, which we set as $\eta \hat{\gamma}_{r,s}^i$, $\eta < 1$. With these assumptions, we can find the new threshold for establishing an infection to be:

$$R_0 = \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$
$$\beta_{su} = \sum_{r=c}^{N} \hat{\gamma}_{r,0}^u, \; \beta_{si} = \sum_{r=0}^{N} \sum_{s=1}^{M} \hat{\gamma}_{r,s}^i$$

Details are given in the Supplementary Material. This expression is similar to the one for free virus transmission in model (2), consisting of two components, one stemming from uDNA replication, the other from iDNA replication. Note that, in contrast to model (2),

unintegrated viral genomes are now a factor in this expression, even when we require multiple unintegrated viral genomes to begin infection. Here, c is the threshold number of unintegrated viral genomes required to start cell-to-cell transmission. The term β_{su} is the sum of the $\hat{\gamma}_{r,s}^{u}$ for synapses where at least the threshold number of viral genomes become unintegrated and none become integrated. The term β_{si} is the sum of the $\hat{\gamma}_{r,s}^{i}$ for synapses where at least one transferred virus becomes an integrated provirus.

The parameters $\hat{\gamma}_{r,s}^{i}$ and $\hat{\gamma}_{r,s}^{u}$ (equal to $\eta\hat{\gamma}_{r,s}^{i}$) in the above expression is a composite of different events, including the probability that r+s viruses are transmitted, and what proportion of these viruses become integrated. To relate $\hat{\gamma}_{r,s}^{i}$ to more tangible parameters, we express $\hat{\gamma}_{r,s}^{i}$ as a function of the probability of integration and distribution of viruses transferred per synapse. There is little known about the factors that determine what proportion of viruses become integrated but if we assume that each virus integrates independently at probability p then $\hat{\gamma}_{r,s}^{i}$ and $\hat{\gamma}_{r,s}^{u}$ can be replaced with more explicit formulas, and the threshold for successful infection is now:

$$R_0=\frac{\lambda}{a_i d}\sum_{r=0}^{N}\sum_{s=1}^{M}\hat{\gamma}^h(1-(1-p)^{r+s})+\frac{\eta\lambda}{a_u d}\sum_{r=c}^{N}\hat{\gamma}^h(1-p)^r$$

Here, $\hat{\gamma_h}$ is the probability that h viruses are successfully transferred by iDNA infected cells, where h=r+s, and η is the reduction in the rate of synapse formation of cells infected with a threshold number of unintegrated viral genomes. We can see from this expression that, if viruses integrate independently, as the average number of viruses transferred per cell increases, the relative contribution by unintegrated viral genomes decreases (Figure 2). Intuitively, if a cell transfers a larger number of viruses per synapse, the chance that at least one virus becomes integrated approaches certainty. Note that there need not be complete independence of integration between viruses for this trend to hold. This result remains robust in the context of different distributions for the probability of transferring h viruses, $\hat{\gamma_h}$ (Figure 2). Little is known about the distribution of the number of viruses transferred per synapse so we explore different commonly occurring distributions (Normal, Uniform, or Poisson). For example, if the number of successfully transmitted viruses follows a Poisson ($\bar{h}$) distribution and each virus independently integrates, $\hat{\gamma_h}$ can be expanded and the R_0 for synaptic transmission is given by:

$$R_0=\frac{\lambda}{a_i d}\sum_{r=0}^{N}\sum_{s=1}^{M}\frac{\overline{h}^{r+s}\mathrm{e}^{-\overline{h}}}{(r+s)!}(1-(1-p)^{r+s})+\frac{\eta\lambda}{a_u d}\sum_{r=c}^{N}\frac{\overline{h}^r\mathrm{e}^{-\overline{h}}}{r!}(1-p)^r$$

According to this expression, the relative contribution of uDNA to the basic reproductive ratio of the virus not only depends on the number of viruses that are transmitted per synapse, but also on other parameters such as the number of uDNA copies required for successful replication, c, and the probability of intergration, p. These dependencies are shown in Figure 3.

Using measured parameters to estimate the relative contribution of uDNA to R_0

With the formulations derived so far, we can attempt to quantify the relative contribution of HIV-1 uDNA to the basic reproductive ratio of the virus, using parameter values as far as they are available. It is important to note that the parameter values used here are based on preliminary data that still need to be investigated in more detail, but they can be used to obtain first insights. One of the most important parameters is the virus output from cells that only contain uDNA compared to that observed in productively infected cells that contain iDNA. Following activation of infected resting T cells, it was found that during a time-span of two days, the amount of virus released from uDNA-infected cells was between 14.8% and 21.5% of that observed with iDNA-infected cells (Trinite et al., 2013). These numbers, however, do not tell the whole story because virus output was not measured during the entire life-span of the cells (which would give the "burst size" of infected cells), but only during a two hour window. Typically, slower virus replication correlates with a longer life-span of infected cells (Blankson et al., 2002; Gilchrist et al., 2004). Hence, when considering the total number viruses released during the life-span of the infected cells, it is likely that the relative contribution from uDNA-infected cells compared to iDNA-infected cells is larger than the percentages quoted above. To be conservative, we will assume that the burst size of cells only containing uDNA is 25% compared to that observed with productively infected cells that contain iDNA (this is also consistent with our current data, in preparation). Another important parameter is the probability that when a cell becomes infected and replicates, virus integration will take place. This was estimated in a separate study, where the extent of fluorescence following single-round infection with reporter viruses indicated whether cells were expressing from uDNA or from iDNA (in preparation). We found that among the viruses that will infect a cell and produce offspring virus, about 50% of the genomes are unintegrated (i.e. *p=0.5*). This is obviously a first estimate and we note that this figure could be revised in future research. At this stage, we do not have any solid evidence against the assumption that a single copy of uDNA can result in successful virus reproduction, so we will assume *c=1*.

Using these parameters, for free virus transmission, the contribution of uDNA replication to R_0 is $100*(1-p)r/(p+(1-p)r)\%$, where r is the ratio $(k_u/a_u)/(k_i/a_i)$. With estimated parameter values, this contribution is found to be 20%.

For synaptic transmission, the parameterized model suggests that uDNA can contribute maximally 20% to the overall value of R_0. As mentioned above, the exact contribution depends on the average number of viruses that are successfully transmitted per synapse, a number that is currently not known. While tens to hundreds of viruses are passed through synapses, it is possible that the majority of these viruses fail to productively infect the cell. For example, the majority of viruses could end up as unintegrated virus that remains in a form in which it is a replicative dead end. These transmissions are not accounted for in our model.

Combining free virus and synaptic transmission

Finally, we can combine the results from free virus and synaptic-only transmission to arrive at the following value for R_0 for a system with both synaptic and cell-free transmission:

$$R_0 = p\frac{\lambda\beta k_i}{a_i u d} + \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

$$\beta_{su} = \sum_{r=c}^{N} \hat{\gamma}_{r,0}^u, \ \beta_{si} = \sum_{r=0}^{N}\sum_{s=1}^{M} \hat{\gamma}_{r,s}^i$$

Note that this expression is simply the sum of the R_0 from the cell-free and synaptic transmission models. Although we can estimate the contribution of unintegrated viral genomes to R_0 for both cell-free and synaptic transmission independently, little is known about the relative importance of synaptic and cell-free transmission at the start of infection.

Discussion

Our experimental data have shown that cells infected with HIV-1 genomes that never integrate can nevertheless successfully replicate the virus under specific conditions (Gelderblom et al., 2008; Trinite et al., 2013). Particularly, this phenomenon is observed if resting cells are infected that are subsequently activated (Trinite et al., 2013). We envision this scenario might be particularly important to the initial establishment of the infection, when the virus is seeking to establish a foothold and existence is precarious. At or near sites of virus entry, infection of resting T cells is likely to be a dominant way for the virus to spread (Haase, 2005; Li et al., 2005; Miller et al., 2005; Reilly et al., 2007; Zhang et al., 1999). Therefore, the question arises to what extent uDNA replication contributes to the basic reproductive ratio of the virus and thus to the ability of the virus to successfully grow from low numbers. Based on mathematical models that were partly parameterized with preliminary data, we made a first attempt to answer this question. According to this analysis, the answer depends on the transmission pathway by which the virus spreads through the target cell population. For free virus transmission, uDNA replication was found to contribute about 20% to R_0, while with synaptic transmission, the percentage depends on the number of viruses that successfully infect a cell through the synaptic pathway, with the maximum possible contribution being around 20%. These numbers obviously depend on the parameter estimates which, as mentioned before, are still preliminary in nature and require further investigation.

This result can be interpreted in a variety of ways. Let us assume that uDNA replication contributes 20% to R_0. According to the latest estimates, the basic reproductive ratio of HIV-1 was found to be around $R_0=8$ (Ribeiro et al., 2010). In this case the predicted R_0 resulting from the replication of only iDNA would be $R_0^{(i)}=6.4$, and the predicted R_0 resulting from uDNA replication would be $R_0^{(u)}=1.6$. On the one hand, uDNA contributes relatively little to R_0. If no uDNA replication occurred at all, the virus would still reliably establish a persistent infection. On the other hand, if all virus replication stemming from integrated DNA was inhibited, the basic reproductive ratio of the virus would still be greater than one, promoting virus persistence, even though the replication kinetics are significantly slower. This has important implications for the use of integrase inhibitors (such as raltegravir) in a prophylactic setting that aims to prevent the infection from becoming established. According to the numbers derived here, integrase inhibitors might be at a disadvantage compared with drugs that intervene at other stages of the virus replication

cycle (entry, reverse transcription, maturation), although integrase inhibitors are not typically administered alone. Our calculations also have implications for the viral response to integrase inhibitors during chronic infection. When virus loads are high during chronic infection, the contribution of uDNA to viral replication is probably negligible because most susceptible cells are activated T cells. When virus load has been reduced to very low levels by drug therapy, however, resting T cells might be important targets for the virus. It is possible that at low virus loads, uDNA replication provides an important mechanism by which the virus can persist at low levels in the face of continuous integrase inhibitor therapy. This is an interesting notion in the light of studies that analyzed virus dynamics during raltegravir intensification of HAART therapy (Buzón et al., 2010; Llibre et al., 2012). It is thought that during traditional HAART therapy (not utilizing integrase inhibitors), a certain amount of low level virus replication continues to occur in a subset of patients, contributing to viral persistence during treatment. When raltegravir is added to the treatment regimen, accumulation of episomal HIV-1 DNA is observed in some patients, indicating recently infected cells[27, 28]. At the same time, however, ultrasensitive virus load measurements in the same patient remained stable despite the addition of raltegravir (Llibre et al., 2012). This could indicate that at those low levels, uDNA replication counters the effect of raltegravir treatment, contributing to the stable maintenance of virus load. It is important to note, however, that the dynamics of residual virus replication are complex, influencing not only virus growth but also CD8 T cell activity (Llibre et al., 2012), which can also affect the observed amount of virus load during treatment. The dynamics of residual virus replication and the effect of raltegravir intensification have also been recently studied mathematically in a different context (Luo et al., 2013).

It has to be pointed out that the basic reproductive ratio of HIV-1 has been measured during the exponential growth phase of the virus population that occurs in the primary phase of the infection. It is unclear, whether this represents the correct replication kinetics in the very early stages of the infection. It is possible that during the observed exponential growth phase, virus replication is faster than during the very early stages of the infection, because mainly activated T cells are the targets of infection. If the basic reproductive ratio in the crucial early phase of the infection is lower, then the R_0 resulting from uDNA replication might be less than one. On the other hand, it has been reported that mathematical models assuming a time delay between infection and virus production give rise to higher estimates for R_0 (Nowak et al., 1997). For SIV infection, the maximal R_0 calculated assuming a time delay of 24 hours was reported to be 68 (Nowak et al., 1997). In this case, uDNA replication alone could surely maintain a persistent infection. More detailed kinetic data, especially about the earliest stages of the infection, are needed to make more definite statements.

As is clear from this discussion, parameter estimates and the interpretation of our calculations remain uncertain due to lack of sufficient data at this stage. This analysis, however, serves as a guide for further experimental work to directly measure crucial parameters that will allow more precise calculations.

Finally, we would like to comment on an aspect of our work that was repeatedly discussed throughout the paper: the multiple infection of cells. In our model, we allowed cells to be simultaneously infected with i unintegrated and j integrated viruses. While multiple

infection is not a central focus of our study, it enters in the context of certain assumptions. First, if a cell contains both uDNA and iDNA in the model, viral replication occurs at a rate driven by the iDNA (Gelderblom et al., 2008). Second, we also considered the possibility that multiple uDNA genomes per cell are required to initiate successful viral replication. There is ample evidence from *in vitro* studies that multiple HIV-1 particles can infect the same cell (Chen et al., 2005; Dang et al., 2004; Del Portillo et al., 2011; Levy et al., 2004). This is further accentuated by work demonstrating that synaptic transmission leads to the simultaneous transfer of tens to hundreds of viruses from the source cell to the target cell (Chen et al., 2007; Del Portillo et al., 2011; Hubner et al., 2009; Jolly et al., 2004; Rudnicka et al., 2009; Sattentau, 2008). The relevance of these processes *in vivo* are a little less clear. Multiple infection was shown to occur during virus growth in animal models (Levy et al., 2004). In a clinical setting, cells from the spleen of infected patients were shown to contain on average 3–4 proviruses (Jung et al., 2002). Other studies, however, argued that most cells in the blood and tissue are singly infected (Josefsson et al., 2013; Josefsson et al., 2011). The reason for this discrepancy is not known, although it could lie in different T cell subsets examined or from more rapid death of cells containing multiple genomes. However, the occurrence of viral recombination, and the likely contribution of this process to viral evolution *in vivo* indicates that multiple infection is also an important driving force *in vivo* (Jung et al., 2002; Levy et al., 2004; Neher and Leitner, 2010).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Anderson, RM.; May, RM. Infectious diseases of humans. Oxfors University Press; Oxofrd, England: 1991.

Blankson JN, Persaud D, Siliciano RF. The challenge of viral reservoirs in HIV-1 infection. Annu Rev Med. 2002; 53:557–93. [PubMed: 11818490]

Bonhoeffer S, Nowak MA. Pre-existence and emergence of drug resistance in HIV-1 infection. Proc R Soc Lond B Biol Sci. 1997; 264:631–7.

Bonhoeffer S, May RM, Shaw GM, Nowak MA. Virus dynamics and drug therapy. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94:6971–6976. [PubMed: 9192676]

Butler SL, Hansen MS, Bushman FD. A quantitative assay for HIV DNA integration in vivo. Nature Medicine. 2001; 7:631–634.

Buzón MJ, Massanella M, Llibre JM, Esteve A, Dahl V, Puertas MC, Gatell JM, Domingo P, Paredes R, Sharkey M. HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. Nature Medicine. 2010; 16:460–465.

Cara A, Cereseto A, Lori F, Reitz MS Jr. HIV-1 protein expression from synthetic circles of DNA mimicking the extrachromosomal forms of viral DNA. J Biol Chem. 1996; 271:5393–7. [PubMed: 8621393]

Chen J, Dang Q, Unutmaz D, Pathak VK, Maldarelli F, Powell D, Hu WS. 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. 2005; 79:4140–9. [PubMed: 15767415]

Chen P, Hubner W, Spinelli MA, Chen BK. Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. J Virol. 2007; 81:12582–95. [PubMed: 17728240]

Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. Proc Natl Acad Sci U S A. 1998; 95:8869–73. [PubMed: 9671771]

Dang Q, Chen J, Unutmaz D, Coffin JM, Pathak VK, Powell D, KewalRamani VN, Maldarelli F, Hu WS. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc Natl Acad Sci U S A. 2004; 101:632–7. [PubMed: 14707263]

Del Portillo A, Tripodi J, Najfeld V, Wodarz D, Levy DN, Chen BK. Multiploid inheritance of HIV-1 during cell-to-cell infection. J Virol. 2011; 85:7169–76. [PubMed: 21543479]

Gelderblom HC, Vatakis DN, Burke SA, Lawrie SD, Bristol GC, Levy DN. Viral complementation allows HIV-1 replication without integration. Retrovirology. 2008; 5:60. [PubMed: 18613957]

Germain RN, Robey EA, Cahalan MD. A decade of imaging cellular motility and interaction dynamics in the immune system. Science. 2012; 336:1676–81. [PubMed: 22745423]

Gilchrist MA, Coombs D, Perelson AS. Optimizing within-host viral fitness: infected cell lifespan and virion production rate. J Theor Biol. 2004; 229:281–8. [PubMed: 15207481]

Haase AT. Perils at mucosal front lines for HIV and SIV and their hosts. Nat Rev Immunol. 2005; 5:783–92. [PubMed: 16200081]

Haase AT. Early events in sexual transmission of HIV and SIV and opportunities for interventions. Annu Rev Med. 2011; 62:127–39. [PubMed: 21054171]

Hatziioannou T, Del Prete GQ, Keele BF, Estes JD, McNatt MW, Bitzegeio J, Raymond A, Rodriguez A, Schmidt F, Mac Trubey C, Smedley J, Piatak M Jr, KewalRamani VN, Lifson JD, Bieniasz PD. HIV-1-induced AIDS in monkeys. Science. 2014; 344:1401–5. [PubMed: 24948736]

Heffernan J, Smith R, Wahl L. Perspectives on the basic reproductive ratio. Journal of The Royal Society Interface. 2005; 2:281–293.

Hladik F, McElrath MJ. Setting the stage: host invasion by HIV. Nat Rev Immunol. 2008; 8:447–57. [PubMed: 18469831]

Hubner W, McNerney GP, Chen P, Dale BM, Gordon RE, Chuang FY, Li XD, Asmuth DM, Huser T, Chen BK. Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science. 2009; 323:1743–7. [PubMed: 19325119]

Jolly C, Kashefi K, Hollinshead M, Sattentau QJ. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. J Exp Med. 2004; 199:283–93. [PubMed: 14734528]

Josefsson L, Palmer S, Faria NR, Lemey P, Casazza J, Ambrozak D, Kearney M, Shao W, Kottilil S, Sneller M, Mellors J, Coffin JM, Maldarelli F. Single cell analysis of lymph node tissue from HIV-1 infected patients reveals that the majority of CD4+ T-cells contain one HIV-1 DNA molecule. PLoS Pathog. 2013; 9:e1003432. [PubMed: 23818847]

Josefsson L, King MS, Makitalo B, Brannstrom J, Shao W, Maldarelli F, Kearney MF, Hu WS, Chen J, Gaines H, Mellors JW, Albert J, Coffin JM, Palmer SE. Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. Proc Natl Acad Sci U S A. 2011; 108:11199–204. [PubMed: 21690402]

Jung A, Maier R, Vartanian JP, Bocharov G, Jung V, Fischer U, Meese E, Wain-Hobson S, Meyerhans A. Multiply infected spleen cells in HIV patients. Nature. 2002; 418:144. [PubMed: 12110879]

Kinter A, Moorthy A, Jackson R, Fauci AS. Productive HIV infection of resting CD4+ T cells: role of lymphoid tissue microenvironment and effect of immunomodulating agents. AIDS Res Hum Retroviruses. 2003; 19:847–56. [PubMed: 14585216]

Komarova NL, Wodarz D. Virus dynamics in the presence of synaptic transmission. Math Biosci. 2013; 242:161–71. [PubMed: 23357287]

Komarova NL, Urwin E, Wodarz D. Accelerated crossing of fitness valleys through division of labor and cheating in asexual populations. Sci Rep. 2012a; 2:917. [PubMed: 23209877]

Komarova NL, Levy DN, Wodarz D. Effect of synaptic transmission on viral fitness in HIV infection. PLoS One. 2012b; 7:e48361. [PubMed: 23166585]

Komarova NL, Levy DN, Wodarz D. Synaptic transmission and the susceptibility of HIV infection to anti-viral drugs. Scientific reports. 2013a:3.

Komarova NL, Anghelina D, Voznesensky I, Trinite B, Levy DN, Wodarz D. Relative contribution of free-virus and synaptic transmission to the spread of HIV-1 through target cell populations. Biol Lett. 2013b; 9:20121049. [PubMed: 23269844]

Levy DN, Aldrovandi GM, Kutsch O, Shaw GM. Dynamics of HIV-1 recombination in its natural target cells. Proc Natl Acad Sci U S A. 2004; 101:4204–9. [PubMed: 15010526]

Li Q, Duan L, Estes JD, Ma ZM, Rourke T, Wang Y, Reilly C, Carlis J, Miller CJ, Haase AT. Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. Nature. 2005; 434:1148–52. [PubMed: 15793562]

Lifson JD, Nowak MA, Goldstein S, Rossio JL, Kinter A, Vasquez G, Wiltrout TA, Brown C, Schneider D, Wahl L, Lloyd AL, Williams J, Elkins WR, Fauci AS, Hirsch VM. The extent of early viral replication is a critical determinant of the natural history of simian immunodeficiency virus infection. J Virol. 1997; 71:9508–14. [PubMed: 9371613]

Little SJ, McLean AR, Spina CA, Richman DD, Havlir DV. Viral dynamics of acute HIV-1 infection. J Exp Med. 1999; 190:841–50. [PubMed: 10499922]

Llibre JM, Buzón MJ, Massanella M, Esteve A, Dahl V, Puertas MC, Domingo P, Gatell JM, Larrouse M, Gutierrez M. Treatment intensification with raltegravir in subjects with sustained HIV-1 viraemia suppression: a randomized 48-week study. Antiviral therapy. 2012; 17:355. [PubMed: 22290239]

Luo R, Cardozo EF, Piovoso MJ, Wu H, Buzon MJ, Martinez-Picado J, Zurakowski R. Modelling HIV-1 2-LTR dynamics following raltegravir intensification. J R Soc Interface. 2013; 10:20130186. [PubMed: 23658114]

Mellors JW, Rinaldo CR Jr, Gupta P, White RM, Todd JA, Kingsley LA. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science. 1996; 272:1167–70. [PubMed: 8638160]

Miller CJ, Li Q, Abel K, Kim EY, Ma ZM, Wietgrefe S, La Franco-Scheuch L, Compton L, Duan L, Shore MD, Zupancic M, Busch M, Carlis J, Wolinsky S, Haase AT. Propagation and dissemination of infection after vaginal transmission of simian immunodeficiency virus. J Virol. 2005; 79:9217–27. [PubMed: 15994816]

Neher RA, Leitner T. Recombination rate and selection strength in HIV intra-patient evolution. PLoS Comput Biol. 2010; 6:e1000660. [PubMed: 20126527]

Nishimura Y, Martin MA. The acute HIV infection: implications for intervention, prevention and development of an effective AIDS vaccine. Curr Opin Virol. 2011; 1:204–10. [PubMed: 21909345]

Nowak, MA.; May, RM. Mathematical principles of immunology and virology. Oxford University Press; 2000. Virus dynamics.

Nowak MA, Lloyd AL, Vasquez GM, Wiltrout TA, Wahl LM, Bischofberger N, Williams J, Kinter A, Fauci AS, Hirsch VM, Lifson JD. Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. J Virol. 1997; 71:7518–25. [PubMed: 9311831]

Perelson AS. Modelling viral and immune system dynamics. Nature Rev Immunol. 2002; 2:28–36. [PubMed: 11905835]

Perelson AS, Ribeiro RM. Modeling the within-host dynamics of HIV infection. BMC Biol. 2013; 11:96. [PubMed: 24020860]

Reilly C, Wietgrefe S, Sedgewick G, Haase A. Determination of simian immunodeficiency virus production by infected activated and resting cells. AIDS. 2007; 21:163–8. [PubMed: 17197806]

Ribeiro RM, Qin L, Chavez LL, Li D, Self SG, Perelson AS. Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. J Virol. 2010; 84:6096–102. [PubMed: 20357090]

Rudnicka D, Feldmann J, Porrot F, Wietgrefe S, Guadagnini S, Prevost MC, Estaquier J, Haase AT, Sol-Foulon N, Schwartz O. Simultaneous cell-to-cell transmission of human immunodeficiency virus to multiple targets through polysynapses. J Virol. 2009; 83:6234–46. [PubMed: 19369333]

Sattentau Q. Avoiding the void: cell-to-cell spread of human viruses. Nat Rev Microbiol. 2008; 6:815–26. [PubMed: 18923409]

Sloan RD, Wainberg MA. The role of unintegrated DNA in HIV infection. Retrovirology. 2011; 8:52. [PubMed: 21722380]

Thalhauser CJ, Lowengrub JS, Stupack D, Komarova NL. Selection in spatial stochastic models of cancer: migration as a key modulator of fitness. Biol Direct. 2010; 5:21. [PubMed: 20406439]

Trinite B, Ohlson EC, Voznesensky I, Rana SP, Chan CN, Mahajan S, Alster J, Burke SA, Wodarz D, Levy DN. An HIV-1 Replication Pathway Utilizing Reverse Transcription Products That Fail To Integrate. J Virol. 2013; 87:12701–20. [PubMed: 24049167]

Wodarz, D. Killer cell dynamics: mathematical and computational approaches to immunology. Springer; New York: 2006.

Wu Y, Marsh JW. Selective transcription and modulation of resting T cell activity by preintegrated HIV DNA. Science. 2001; 293:1503–6. [PubMed: 11520990]

Zhang Z, Schuler T, Zupancic M, Wietgrefe S, Staskus KA, Reimann KA, Reinhart TA, Rogan M, Cavert W, Miller CJ, Veazey RS, Notermans D, Little S, Danner SA, Richman DD, Havlir D, Wong J, Jordan HL, Schacker TW, Racz P, Tenner Racz K, Letvin NL, Wolinsky S, Haase AT. Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. Science. 1999; 286:1353–7. [PubMed: 10558989]

Zhang ZQ, Wietgrefe SW, Li Q, Shore MD, Duan L, Reilly C, Lifson JD, Haase AT. Roles of substrate availability and infection of resting and activated CD4+ T cells in transmission and acute simian immunodeficiency virus infection. Proc Natl Acad Sci U S A. 2004; 101:5640–5. [PubMed: 15064398]
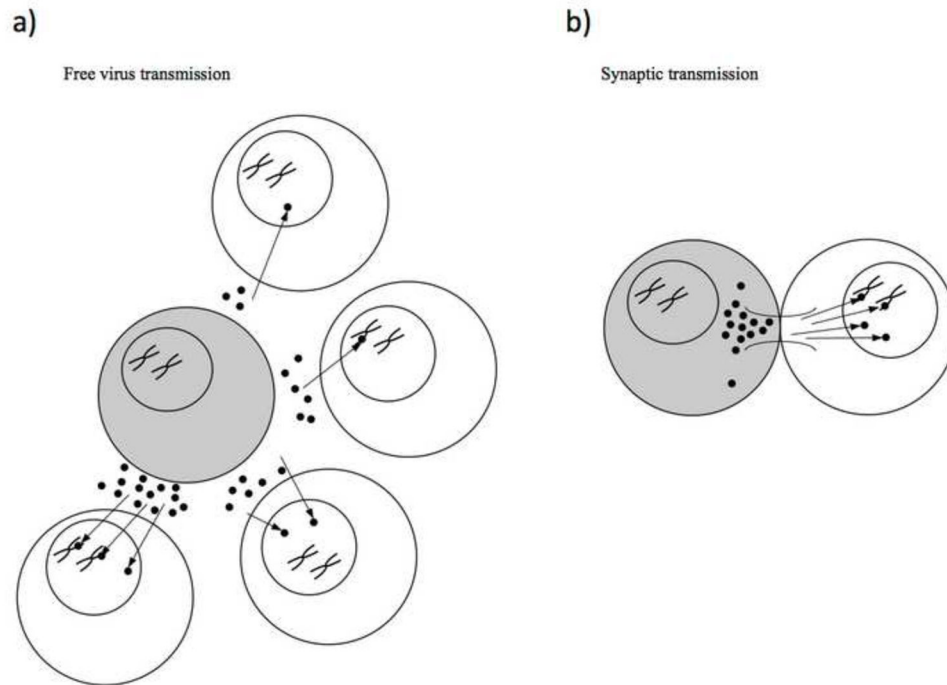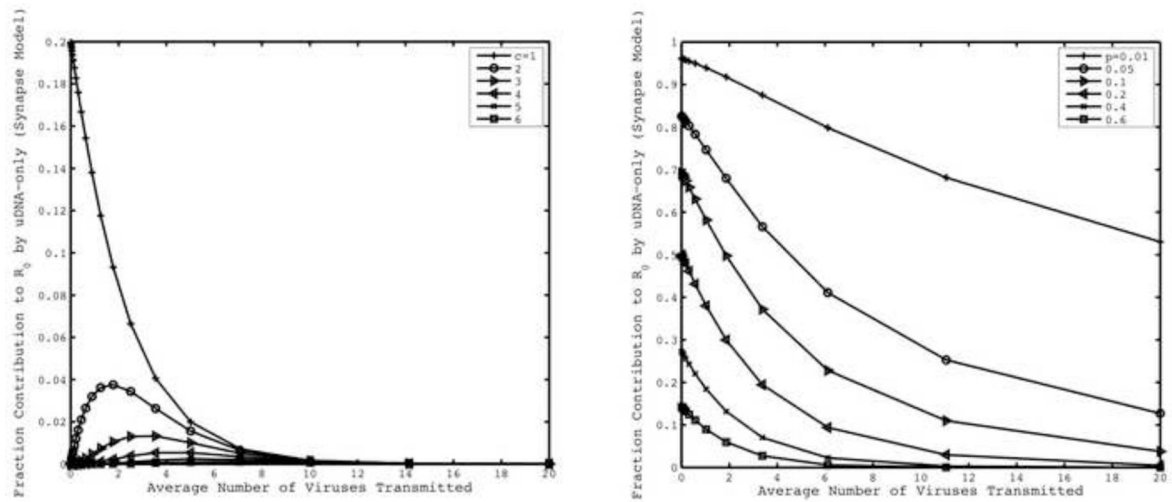
Figure 1.
Schematic of free virus and synaptic transmission modes. (a) In free virus transmission, viruses are released by an infected cell (grey) and can enter the target cell population. Different virus fates are shown with different target cells. Clockwise from the top: a single virus infects the target cell but remains unintegrated, a single virus infects a target cell and is integrated into the genome, multiple viruses infect a target cell but all remain unintegrated, and multiple viruses infect a target cell and some become integrated. (b) Synaptic transmission potentially involves a high number of viruses transferring into a target cell at once. These viruses can be integrated into the genome or remain unintegrated. In contrast to free virus transmission, the multiplicity of infection can be much higher. A single infected cell can potentially transmit via both modes of transmission.
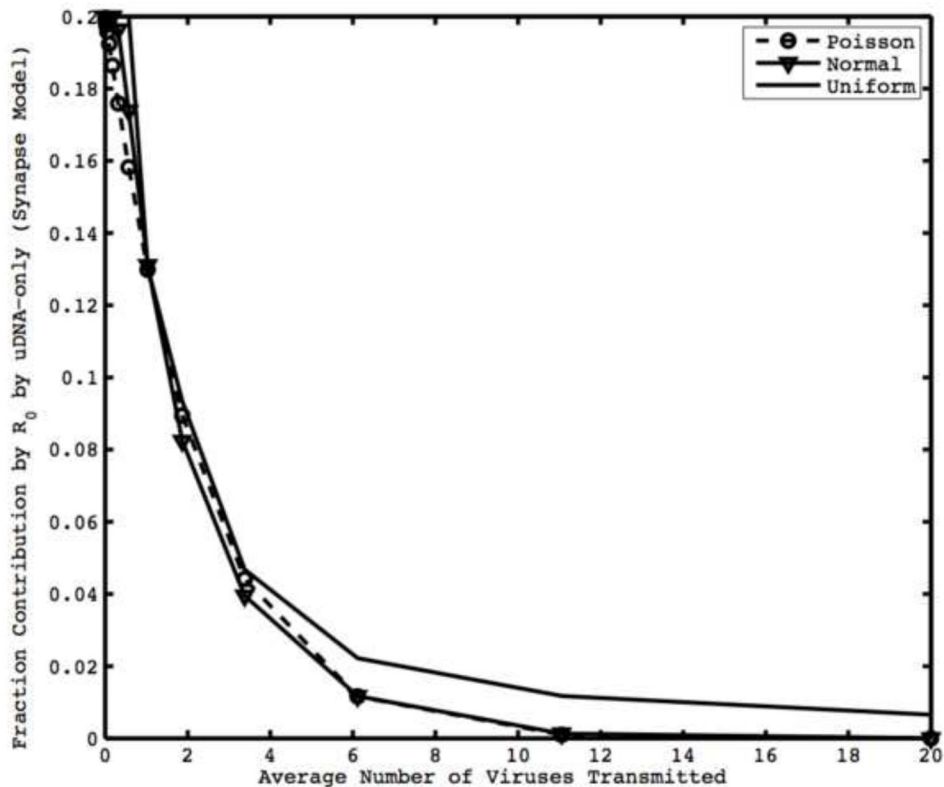
Figure 2.
Fraction contribution to R_0 by cells with uDNA-only for synaptic transmission. Lines are different distributions of the number of viruses transferred per synapse. The contribution to R_0 by uDNA depends on average number of viruses transmitted but remains relatively constant with respect to the distribution of viruses transmitted. We assume that viruses integrate independently. Parameters are $\lambda = 10000$, $d=0.01$, $\eta=0.25$, $p=0.5$, $c=1$, $a_i=0.7$, $a_u=0.7$.

Figure 3.
(left) Contribution of uDNA towards R_0 for synaptic transmission with varying threshold number of uDNA copies required for productive infection, c. (right) Contribution of uDNA towards R_0 for synaptic transmission with varying integration probabilities, p. We assume that viruses integrate independently and number of viruses transmitted follows a Poisson distribution. Default parameters are $\lambda = 10000$, $d=0.01$, $\eta=0.25$, $p=0.5$, $c=1$, $a_i=0.7$, $a_u=0.7$.