UNIVERSITY OF CALIFORNIA,
IRVINE

Population Dynamics of Human Immunodeficiency Virus

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTORATE OF PHILOSOPHY

in Ecology and Evolutionary Biology

by

John Wei Lau

Dissertation Committee:
Professor Dominik Wodarz, Chair
Associate Professor Matthew McHenry
Associate Professor Kevin Thornton

2014

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my committee chair and advisor Dominik Wodarz for his patience and advice over the years. Without his help this dissertation would not have been possible.

I would also like to thank Associate Professors Matthew McHenry and Kevin Thornton for their service on my committee.

I thank the Department of Ecology and Evolutionary Biology at UC Irvine for support. I also thank the Mathematical, Computational, and Systems Biology program at UCI.

Financial support was provided by the Department of Ecology and Evolutionary Biology, the Mathematical, Computational, and Systems Biology program, and NIAID grant 5R01AI093998-04.

CURRICULUM VITAE

John Wei Lau

2008	B.A. in Applied Mathematics, University of California, Davis
2008	B.S. in Biochemistry and Molecular Biology, University of California, Davis
2009-10	Gateway Program in Mathematical, Computational, and Systems Biology. University of California, Irvine
2010-14	Graduate Student Researcher, Ecology and Evolutionary Biology, University of California, Irvine
2010-14	Teaching Assistant, Department of Ecology of Evolutionary Biology, University of California, Irvine
2014	M.S. in Statistics, University of California, Irvine
2014	Ph.D. in Ecology and Evolutionary Biology, University of California, Irvine

FIELD OF STUDY

Population Dynamics

PUBLICATIONS

Lau, John W; Levy, David N. Wodarz, Dominik. "Contribution of HIV-1 genomes that do not integrate to the basic reproductive ratio of the virus". Journal of Theoretical Biology. Accepted. December 2014.

Wodarz, Dominik; Sun, Z; Lau, JW; Komarova, NL. "Nearest-neighbor interactions, habitat fragmentation, and the persistence of host-pathogen systems" Am Nat. 2013 Sep;182(3):E94-E111. doi: 10.1086/671185. Epub 2013 Jul 18.

Wodarz, Dominik; Hofacre, Andrew; Lau, John W; Sun, Zhiying; Fan, Hung; Komarova, Natalia L. "Complex spatial dynamics of oncolytic viruses in vitro: mathematical and experimental approaches." PLoS Computational Biology. 2012: 8 (6):e1002547

ABSTRACT OF DISSERTATION

Population Dynamics of Human Immunodeficiency Virus

By

John Wei Lau

Doctor of Philosophy in Ecology and Evolutionary Biology

University of California, Irvine, 2014

Professor Dominik Wodarz, Chair

In this work, we explore the *in vivo* population dynamics of HIV with mathematical and computational methods. Chapter 1 examines the effects of productive unintegrated virus (uDNA) on viral dynamics in the context of free-virus and synaptic transmission. We find that productive uDNA can contribute up to 20% towards the basic reproductive ratio of HIV *in vivo*. If more than one unintegrated virus is required for productive infection, then uDNA does not contribute towards R_0 from free virus infection. As more viruses are successfully transmitted per synapse, the lower the contribution of uDNA. Chapter 2 explores the effect of uDNA in the context of an immune response during the asymptomatic phase. We find that productive uDNA can decrease or increase set-point viral levels compared to the case of inert uDNA, and this depends on the rate of viral production and the strength of the immune response. In Chapter 3, we explore a two-compartment model to explain the observed difference in the multiplicity of infection of HIV between the lymphoid system and the blood. Our model suggests that the absence of strong synaptic transmission in the peripheral blood tends to create many singly infected cells – reducing the proportion of multiply infected cells. Although this simple mechanism can explain much

of the difference between the two systems, we also conclude that there must be some additional mechanism that is reducing the number of highly multiply infected cells in the blood.

INTRODUCTION

The principles and methods used in the study of population dynamics can be applied to any collection of units and their interaction with other factors. In the case of this dissertation, we investigate the ecology of Human Immunodeficiency Virus (HIV) *in vivo* with mathematical and computational models. We treat HIV and susceptible cells as individual organisms that interact with each other and the environment within an infected individual. HIV was chosen in particular because of the rich history of mathematical modeling of the disease and availability of data to support and validate models. New aspects of the disease have been discovered recently and, in this dissertation, I explore different facets of HIV from the perspective of population dynamics.

Human Immunodeficiency Virus (HIV) is a retrovirus that causes acquired immunodeficiency syndrome (AIDS). As of 2012, approximately 35 million people are living with HIV worldwide, with 2.5 million infections, and 1.6 million deaths [1]. A majority of the new infections are in Sub-Saharan Africa, where education and medicine is less available [1, 2]. In well-developed countries the number of AIDS deaths and new HIV cases has declined, partially due to the availability of effective treatment [1]. However, even amongst those with access to medicine, the disease remains chronic and requires a life-long drug regimen. Despite significant advances in treatment, a permanent cure to HIV remains elusive.

HIV and AIDS gained prominence in the early 1980s. It was during this time that the United States Centers for Disease Control coined the name AIDS and began dedicating groups to study and combat the potential epidemic. It is believed that HIV was originally transferred from non-human primates, who have the similar disease SIV (simian immunodeficiency virus), through blood contact. After its recognition in the United States, in 1981, there were approximately 600 recorded deaths from AIDS in 1982, ballooning to 200,000 cumulative deaths in 1992 and 500,000 by 2002 [3]. In 2002, AIDS was the leading cause of premature death worldwide amongst people from 15-59 [2]. As antiretroviral drugs that were developed in the 1990s became more available, new infections in developed nations declined [1]. However, worldwide, particularly in Sub-Saharan Africa, the disease remained prevalent and the number of new infections has risen [1,2].

Several anti-viral drugs are currently available but the treatments are chronic, expensive, and are often not readily available in lesser-developed countries [3]. Drug treatments usually involve a cocktail of 3 or more drugs that target specific proteins or stages of the viral life cycle. For example, the first such drug, azidothymidine (AZT), is an inhibitor that targets the viral protein reverse transcriptase. As a cocktail of multiple drugs, which has been reduced to a single pill, these drugs have advanced to reduce serum virus to low or undetectable levels [4]. In addition, treatment can be used as a preventative tool, and those under a drug regimen against HIV have a significantly reduced chance of transmitting the disease [5]. Despite significant advances in drug treatment, invariably, on cessation of treatment, the virus resurges, and the patient must continue to take the drugs for life [6, 7].

The progression of HIV in a patient is variable but the stages of the disease are well known. The initial phase presents with mild symptoms and patients are considered to be the acute phase that is characterized by high viral loads in the blood. The acute phase typically lasts for weeks until patients enter an asymptomatic phase that typically last 5-10 years in a majority of patients. During this period, serum virus levels are lower compared to the acute phase. However, the virus is not inactive during this time – it is constantly replicating and the immune system is constantly combatting it, as shown by the rapid viral turnover during the asymptomatic phase [8, 9]. Since the asymptomatic phase is measured in years, it can be considered a dynamic steady state for the purpose of mathematical modeling. After the asymptomatic phase, if left untreated, patients enter the AIDS phase of the disease. The AIDS phase is characterized by a reduction of CD4+ T-cell levels below 200 cells per microliter or the presence of characteristic opportunistic infections. At this point, bacterial, fungal, or viral infections that are normally under control by the immune system can develop and these diseases ultimately kill the infected individual.

HIV primarily targets immune cells including CD4+ T lymphocytes, dendritic cells, and macrophages, ultimately weakening the immune system and allowing for opportunistic infections. The virus uses the CD4 receptor found on these cells to enter and begin replication. Susceptible cells have different half-lives: productively infected cells have half-lives on the order of hours or days while latently infected memory CD4+ T cells can survive for months or years [10-13]. However, CD4+ T lymphocytes are the largest population of cells that are susceptible to the disease and are the focus of this dissertation.

The HIV-1 virion includes two copies of single-stranded RNA, together with key proteins inside a viral envelope. Among the proteins are integrase, reverse transcriptase, protease, and others that enhance viral replication through interactions with the target cell. Reverse transcriptase is responsible for transcribing RNA to DNA inside the nucleus. The reverse transcriptase enzyme can switch between available transcripts, creating new permutations that may have a selective advantage in the presence of an immune response or drug regimen [14]. The intregrase enzyme is responsible for integrating the newly created viral DNA transcript into the host genome. Once a viable copy of HIV is inserted into the genome, it can be expressed with the host machinery to create new viral transcripts.

The molecular biology of infection of HIV-1 is complicated and there has been much research devoted into elucidating the proteins and pathways involved in transmission [15]. For the purpose of this dissertation, a simplified model of HIV transmission is used that captures the relevant details without needless complication. In free virus transmission, the virus binds to a cell with an available CD4+ receptor and co-receptor such as CCR5 or CXCR4 [16, 17]. The contents of the viral envelope are released, including viral RNA and the enzymes necessary to begin production of progeny virus. Viral RNA is reverse transcribed into DNA, and forms a pre-integration complex with viral proteins [18]. The pre-integration complex enters the nucleus and the viral DNA is then integrated into the host genome via viral integrase. In productive infection, the integrated virus is expressed and produces progeny viral RNA which is translated into viral proteins. Viral RNA and proteins are then

packaged into new virions that can be released from the infected cell. The entire process can occur in the order of days but the provirus can also remain unexpressed and dormant within the cell as long as the cell remains alive.

Although many HIV+ patients have the disease managed, a permanent and universal cure for HIV remains elusive due to a variety of factors inherent to the disease. The first is the high rate of mutation due to the recombinatorial and error-prone nature of reverse transcriptase [19]. Reverse transcriptase has no error-correcting mechanism and mutations can occur during the transcription process. More importantly, reverse transcriptase can switch between different viral templates, amplifying the diversity of HIV. The virus evades the adaptive immune response by constantly mutating within a patient. These mutations can cause the virus to replicate faster or use different co-receptors [20, 21]. Secondly, HIV remains latent in the infected individual and, invariably, these latent viruses resurface after drug therapy is stopped [23-24]. Latent viruses include unexpressed provirus that replicates as the infected cell replicates, remaining in perpetuity for as long as the cell line continues to exist [25]. In addition, anatomical reservoirs of immune cells can be untouched by drug therapy and can persist during the chronic phase of the disease.

This dissertation explores the in-host population dynamics of HIV, particularly in the context of three relatively recently discovered characteristics – multiple infection of HIV, viral spread via synaptic transmission, and the activity of unintegrated viral genomes. It has been shown that a single cell can be infected with multiple copies of HIV. This has been shown to occur *in vitro* but there have been differing reports on the prevalence of *in*

vivo multiple infection [19, 26-28]. Some reports suggest that multiple infection is common [29-32] while others suggest that multiple infection is rare [33, 34]. Multiple infection by free virus transmission must occur prior to the down-modulation of surface receptors following initial infection of the cell [35]. During this window, multiple HIV viruses can potentially infect the cell. These viruses, if genetically distinct due to previous recombination or mutation, can potentially recombine inside a single cell and increase overall genetic diversity [36]. The presence of multiple viruses competing for the same cellular resources can potentially have an effect on viral dynamics. In addition, multiple viruses can potentially be detrimental to each virus due to increased cellular or immune responses. If different viral epitopes, from different quasispecies of the virus, are presented on the cell surface, then a more efficient immune response may be mounted, which can have a detrimental effect on each virus inside the cell. The dynamics of multiple infections can be complex and has not been explored in detail.

In addition to free virus transmission, HIV can also transmit directly to other susceptible cells via a viral synapse [37-41]. The exact mechanism is unclear, but is believed to involve the same receptors that are used in T-cell communication [38]. This process is clearly density dependent, and is therefore not likely to occur in regions where CD4+ cells are sparse. The infected cell can transmit hundreds of immature virus particles to the target cell. It is unclear how many of these viruses are viable but due to the number of viruses transferred per synapse it may be a significant mode of transmission. The relative strength of synaptic and free virus transmission is unknown, may depend on individual and viral factors, and may change as the disease progresses.

It was recently shown that unintegrated viruses can potentially produce progeny viruses [42, 43]. Integrase negative mutants can still produce viruses, albeit at a slower rate [43]. However, the cells infected with these viruses have a longer lifespan and may be more efficient at evading the immune response due to their slower virus production rate. These unintegrated viruses exist as stable episomes within the nucleus [44]. These episomal viruses can be transcribed and new virions can be produced. In addition, it would not be surprising that, in the presence of productive integrated virus, the episomal RNA is more likely to be packaged into a virion [42]. The unintegrated viruses can potentially employ the machinery used by the integrated viruses and can add to viral diversity.

In general, the mathematical modeling of population dynamics has provided insights into complicated biological and ecological systems [45-51]. In a system with many factors that interact in non-linear ways, it may be difficult to understand and pinpoint the key components. Mathematical models simplify the complex biological systems and allow us to present reasonable hypotheses based on these models. Once formulated, we can adjust the model to explore the effects of modifying the system in specific ways. Mathematical models also allow us to estimate key parameters like growth and death rates, which are necessary to understanding the dynamics of any population, including the progression of HIV [8-9, 11, 52].

Mathematical modeling has specifically been used successfully to model the *in vivo* population dynamics of HIV. These models have allowed us to identify key characteristics

of the virus, including lifespan, rate of replication, and rate of clearance of the virus [45-49, 52]. Estimating these parameters is essential to understanding the progression and evolution of the virus in the host. Models have also given us insight to viral evolution, immune escape, and the emergence of drug-resistant forms [49,53-55]. In addition, mathematical modeling has been invaluable in the formulation of new hypothesis and as a guide for clinical and laboratory experiments [49].

Each chapter in this dissertation explores a different aspect of HIV through mathematical modeling and analysis. Chapter 1 focuses on the effect of unintegrated genomes on viral dynamics at the start of infection. Chapter 2 expands on this by examining unintegrated genomes during the dynamic asymptomatic phase and in the context of an immune response. Finally, chapter 3 explores a series of two-compartment models to attempt to explain the difference in multiplicity of infection between the peripheral blood and lymphoid system.

CHAPTER 1: Effect of HIV replication from unintegrated genomes

On basic virus dynamics

Abstract

Recent experimental data has shown that T-cells with unintegrated viral DNA (uDNA) can successfully produce virus. We use mathematical models to estimate the effect of these unintegrated viruses on the *in vivo* basic reproductive ratio. Our model includes synaptic and free-virus transmission, as well as multiple infection of target cells. We find that, with preliminary parameter estimates, uDNA can contribute up to 20% to the total reproductive ratio. If more than one uDNA is required for productive infection, then uDNA does not contribute to R_0 for free virus transmission. For synaptic transmission, uDNA contributes to R_0 regardless of the number of uDNA required for replication. The more viruses transferred per synapse, the lower the contribution of uDNA because this increases the chance that at least one virus integrates. Similarly, the higher the probability of integration, the lower the effect of uDNA. If uDNA contributes 20% to the total R_0, then we estimate that R_0=1.6 from uDNA only.. This would suggest that viral production from cells infected with uDNA alone is enough to establish infection. This has potential implications on the effectiveness of integrase inhibitors as a prophylactic treatment. This work also serves as a mathematical framework for the measurement of critical parameters necessary to understand importance of uDNA in HIV progression.

Introduction

Human immunodeficiency virus-1 (HIV-1) infects cells of the immune system, mainly CD4+ T cells and antigen presenting cells, such as dendritic cells and macrophages. 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 and translation, 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 extensive growth of the HIV population *in vivo* during the acute phase of the infection before the virus population settles around a post-acute steady state, the level of which can be indicative of speed with which the infection progresses from the asymptomatic phase to AIDS.

Extensive clinical and experimental data exist that document these dynamics both in vivo and during in vitro experiments [46-49]. Mathematical models have been very useful to help interpret those data, to measure crucial parameters, and to generate hypotheses. 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 at the beginning of the infection [56]. It has been estimated for HIV in a variety of settings, and has important implications for understanding disease progression as well as the response to anti-viral drug therapy [57].

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). According to the literature, while viral uDNA can be expressed to a certain degree, it is a replicative dead end and does not lead to the production of infectious offspring virus. Recent data, however, indicate that the situation could be more complex [42-43, 58]. If a cell is co-infected with integrated and unintegrated virus, viral uDNA has been shown to successfully produce infectious offspring virus, suggesting cooperative interactions [43]. Even more striking, a recent study has shown that cells that only contain viral uDNA can successfully produce infectious offspring virus. They do so at a rate that is small compared to the rate of virus production from integrated genomes. At the same time, however, cells infected only with uDNA also live significantly longer than iDNA-infected cells, thus increasing the total amount of virus produced by those cells during their life span [43]. Successful uDNA replication, however, does not seem to occur under all conditions. It is observed most readily in resting cells that become infected and subsequently activated, which could be especially relevant to the initial stages of the infection following virus transmission [43].

For the first time we present mathematical model that accounts for successful virus replication both from cells containing integrated virus and from cells that only contain unintegrated virus. The aim 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 the rate of death in the respective cells. In our model, a cell with a single productive 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. 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 the generation of multiply infected cells is a likely event in the context of synaptic transmission. 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.

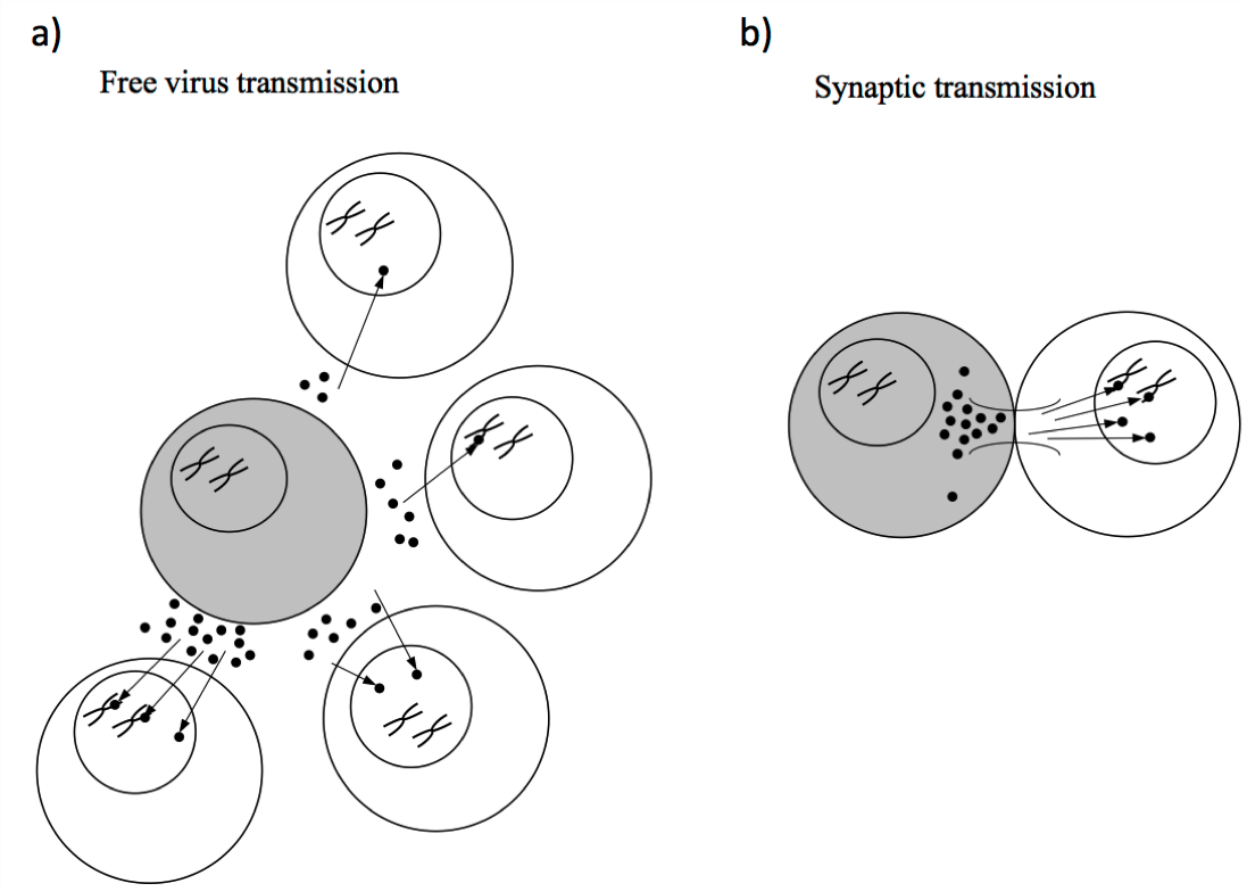


Figure 1.1: Diagram of HIV models 1 and 2. Model 1 has cell-free transmission and model 2 has both cell-free and cell-to-cell transmission. In Model 1, CD4+ T-cells (represented by larger circles) with i unintegrated and j integrated viruses are infected and gain an unintegrated (top cell) or integrated (bottom cell) virus. In Model 2, cell-free infection occurs, but now two cells, one with i unintegrated and j integrated viruses and another cell with k unintegrated and l integrated viruses can interact. The source cell transmits r integrated and s unintegrated viruses, creating a cell with i + r integrated and j + s unintegrated viruses.

The basic model of virus dynamics with free virus transmission

First, we briefly review 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 [45-48]. 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 (Eq 1.1).

(Eq. 1.1)
$$\frac{dx}{dt} = \lambda - dx - \beta xv$$
$$\frac{dy}{dt} = \beta xv - ay$$
$$\frac{dv}{dt} = ky - uv$$

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 and is cleared with a rate u. This model is characterized by two equilibriums. 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 equilibrium. 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 [45] and references therein and table 1.1.

Parameter	Units	Notes
N (max unint)	none	Maximum unintegrated per cell
M (max int)	none	Maximum integrated per cell
λ	Cells $ml^{-1}day^{-1}$	Uninfected production rate
d	day^{-1}	Uninfected death rate
a_i	day^{-1}	Provirus infected death rate without / with immune response
a_u	day^{-1}	Unintegrated infected death rate
k_i	day^{-1}	Prod. rate of free virus from provirus infected cells
k_u	day^{-1}	Prod. rate of free virus from unintegrated infected cells
u	day^{-1}	Clearance rate of free virus
γ	ml $Cells^{-1}$ day^{-1}	Base infectivity for synaptic transmission
β	ml $Cells^{-1}$ day^{-1}	Infectivity for free virus transmission
h	none	Average cells successfully transferred per synapse
p	none	Chance of integration
c	none	Number of unintegrated viruses required for productivity

Table 1.1 : Parameter and definitions used in Eq. 1.1-1.3.

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 cells infected 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.

$$\frac{dy_{0,0}}{dt} = l - dy_{0,0} - \beta y_{0,0} v$$

$$\text{(Eq. 1.2)}\quad \frac{dy_{i,j}}{dt} = (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j}$$

$$\frac{dv}{dt} = \sum_{i=c}^{N} \sum_{j=0}^{M} k_{ij} y_{ij} - uv$$

As in the previous model (Eq. 1.1), infection occurs with a rate . Upon infection, integration of the viral genome occurs with a probability p, while uDNA capable of replication is generated with a probability 1-p. Note that not all types of uDNA is capable of replication [44] and that the uDNA taken into account in the model only includes viral genomes that can replicate. 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. 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).

As in the basic model of virus dynamics, this system is also characterized by two equilibriums. In the absence of infection, the system converges to the trivial equilibrium, given by is $y_{00}= \lambda/d$, $y_{ij}=0$, and $v=0$. If the virus establishes a persistent infection, the system converges to the internal equilibrium where all populations are greater than zero. In the current model, the equilibrium expressions are difficult to obtain and are thus not written out here.

Using standard methods (see Appendix for details), we can calculate the threshold for a successful infection from (Eq 1.2), defined by the stability of the virus-free steady state. This gives rise to the basic reproductive ratio of the virus in model (Eq. 1.2), 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 (Eq. 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 (Eq. 1.1). 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 it 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 further expanded to include not only free virus transmission, but also synaptic transmission of the virus. The model is given by the following set of ordinary differential equations, which are based on previous work [59-62]

$$\begin{aligned}
\frac{dy_{0,0}}{dt} &= l - dy_{0,0} - \beta y_{0,0} v - y_{0,0} \sum_{k=0}^{N} \sum_{l=0}^{M} y_{kl} \sum_{r=0}^{N} \sum_{s=0}^{M} g_{rs}^{kl} \\
\frac{dy_{i,j}}{dt} &= (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j} + \sum_{k=0}^{N} \sum_{l=0}^{M} y_{kl} \left(\sum_{r=0}^{i} \sum_{s=0}^{j} g_{rs}^{kl} y_{i-r,j-s} - y_{ij} \sum_{r=0}^{N-i} \sum_{s=0}^{M-j} g_{rs}^{kl} \right) \\
\frac{dv}{dt} &= \sum_{i=0}^{N} \sum_{j=0}^{M} k_{ij} y_{ij} - uv
\end{aligned} \qquad \text{(Eq1.3)}$$

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

This model with synapsing is modified from a model by [32], which did not consider unintegrated viruses. Compared to model (Eq. 1.2), it has one additional parameter: γ_{rs}^{kl}, which describes the probability that a cell with k unintegrated and l integrated viruses transmits r unintegrated and s integrated viruses. 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 viruses are unintegrated while s viruses become integrated provirus.

The equation describing the dynamics of uninfected cells, x, is identical to model (Eq. 1.2) 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 viruses transferring r unintegrated and s integrated viruses. The equation for infected cells is also identical to that used in model (Eq. 1.2) 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 $i+j \leq M$. For simplicity, we did not explicitly write the equations for the endpoints, i.e. y_{iM}, $y_{Nj,}$ y_{NM}. The equations for these species do not have a sink term from the addition of viruses via free virus transmission or cell-to-cell transmission.

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, γ_{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 viruses, 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 γ_{rs}^{kl} can be simplified to $\hat{g}_{r,s}^{u}$ and $\hat{g}_{r,s}^{i}$, for any k and l. Here, $\hat{g}_{r,s}^{u}$ is the reduced transmission rate from cells with only unintegrated viruses, which we set as $h\hat{g}_{r,s}^{i}$, η < 1.With these assumptions, we can find the new threshold for establishing an infection to be:

$$R_0 = \frac{b_{si} I}{a_i d} + \frac{b_{su} I}{a_u d}$$

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

Details are given in the Appendix. This expression is similar to the one for free virus transmission in model (Eq. 1.2), consisting of two components, one stemming from uDNA replication, the other from iDNA replication. Note that, in contrast to model (Eq. 1.2), unintegrated viruses 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 viruses required to start cell-to-cell transmission. The term β_{su} is the sum of

the $\hat{g}^{u}_{r,s}$ for synapses where at least the threshold number of viruses become unintegrated and none become integrated. The term β_{si} is the sum of the $\hat{g}^{i}_{r,s}$ for synapses where at least one transferred virus becomes an integrated provirus.

The parameters $\hat{g}^{i}_{r,s}$ and $\hat{g}^{u}_{r,s}$ (equal to $h\hat{g}^{i}_{r,s}$) 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 $\hat{g}^{i}_{r,s}$ to more tangible parameters, we express $\hat{g}^{i}_{r,s}$ 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 $\hat{g}^{i}_{r,s}$ and $\hat{g}^{u}_{r,s}$ can be replaced with more explicit formulas, and the threshold for successful infection is now:

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

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 viruses. 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 viruses decreases. 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.

The last unaddressed variable in the above equation is how $\hat{\gamma}_h$, the probability of transferring h viruses, is distributed. 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:

$$R_0 = \frac{l}{a_i d}\sum_{r=0}^{N}\sum_{s=1}^{M}\frac{\bar{h}^{r+s}\exp^{-\bar{h}}}{(r+s)!}(1-(1-p)^{r+s}) + \frac{hl}{a_u d}\sum_{r=c}^{N}\frac{\bar{h}^{r}\exp^{-\bar{h}}}{r!}(1-p)^{r}$$

With this formulation, we can arrive at a first order estimate of the effect of unintegrated viruses on establishing a successful infection for synaptic transmission under the above assumptions (See figure 1.3). With estimated the parameter values (c=1, p=0.5, η=0.25), a maximum of 50% of R_0 is due to unintegrated viruses. In contrast, with only free virus transmission, the contribution to R_0 by uDNA 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, the contribution to R_0 by uDNA is 20%. Roughly, for synaptic transfer, if the R_0 of HIV is approximately 7 new infections per infected cell less than 3.5 of these new infections are due to cells that are only infected with unintegrated viruses [57]. As the threshold for the number of unintegrated viruses required for infected cells to begin producing viruses (the parameter c) increases, the contribution to R_0 by uDNA decreases. As the probability of integration (p) increases, the contribution by uDNA decreases.

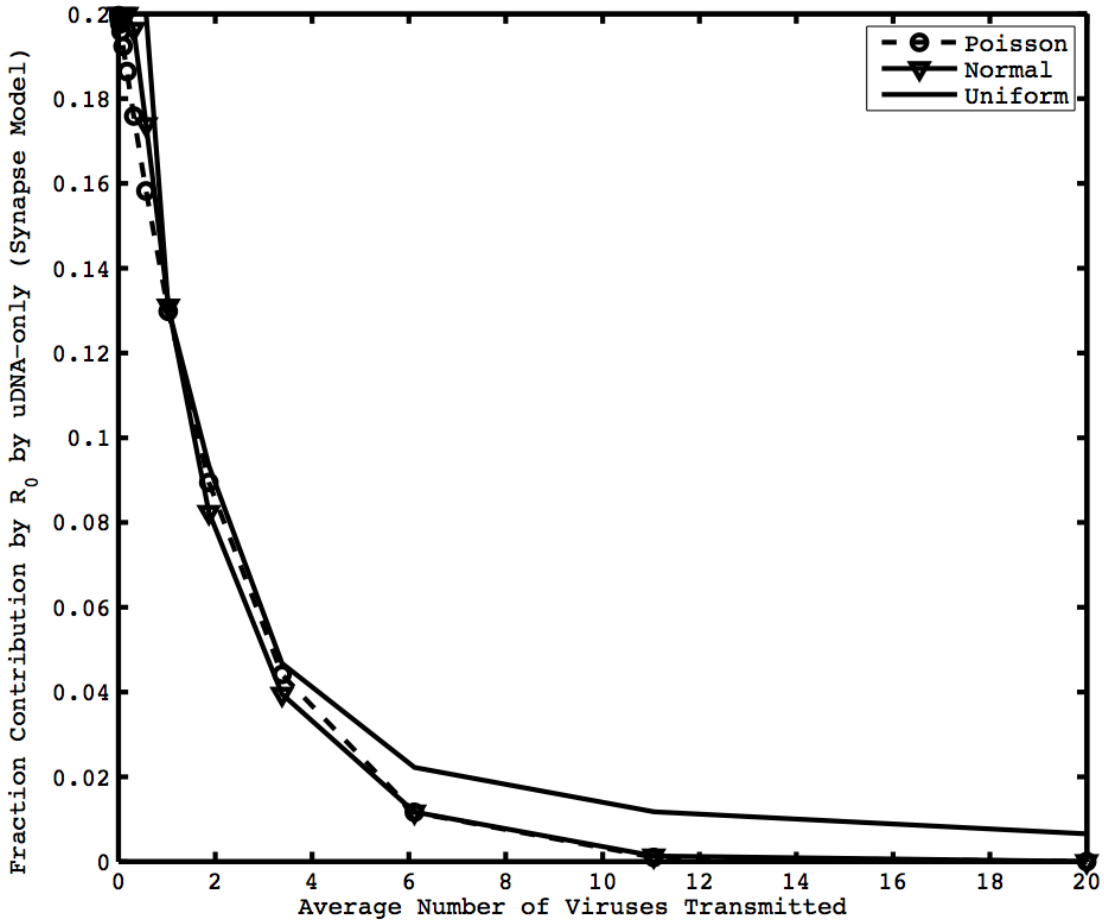


Figure 1.2: Fraction contribution to R_0 by cells with uDNA-only for synaptic transmission. Lines are different distributions of the number of viruses transferred per synapse. The contribution to R_0 by uDNA depends on average number of viruses transmitted but remains relatively constant with respect to the distribution of viruses transmitted. We assume that viruses integrate independently. Parameters are η=0.25, p=0.5, c=1, a_u=0.25a_i.

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:

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

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

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 viruses 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. In our models, if cell-free transmission is the significant mode of transmission for HIV-1 at the beginning of infection, then cells with only unintegrated virus do not contribute significantly towards virus production at the start of HIV-1 infection. If synaptic transmission is the significant mode of transmission then unintegrated viral DNA has a

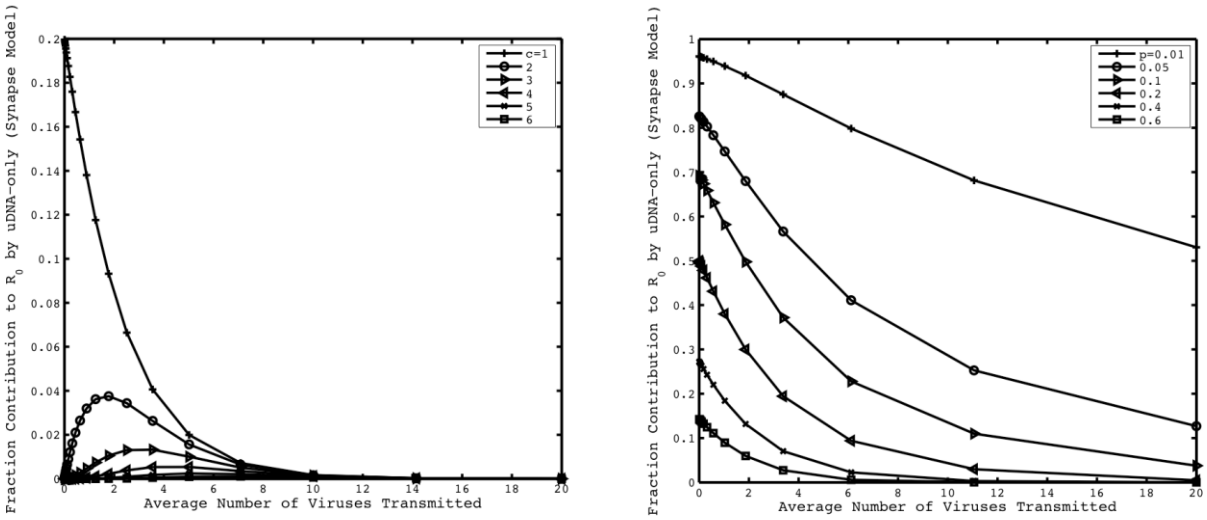


Figure 1.3 left) Contribution of uDNA towards R_0 of synaptic transmission with varying threshold of uDNA required for productive infection. (Right) Contribution of uDNA towards R_0 of synaptic transmission versus the average number of viruses transferred successfully per synapse. As the average number of viruses transmitted per synapse increases, the probability that all transmitted viruses remain unintegrated approaches zero and the contribution of unintegrated viruses decreases. We assume that viruses integrated independently and number of viruses transmitted follows a Poisson distribution. Default parameters are η=0.25, p=0.5, c=1, a_u=0.25a_i.

larger, but still relatively small effect at the start of HIV infection.

Discussion

Recent experimental studies have shown that viral DNA can be expressed even without integrating into the host genome. This is a potential source of genetic diversity and can have an effect on *in vivo* dynamics. Although unintegrated viruses are less productive than integrated virus, they can contribute towards establishing infection. Based on the models we presented here, unintegrated viruses can contribute towards the *in vivo* R_0 of the virus. Under free virus transmission, we found that, with our best estimates of parameter values, replication from uDNA can contribute about 20% to R_0. Under synaptic transmission, the contribution to the reproductive ratio depends on average number of viruses transferred per synapse and probability of integration.

The models used in this paper can serve to estimate critical parameters. It is clear from the analysis that the average number of viruses successfully transferred and integration probability are critical parameters to understand the effect of unintegrated virus. In addition, the relative strengths of synaptic and free-virus transmission are important to understanding the effect of unintegrated virus. If the probability of successful integration is low, then unintegrated virus will have a larger effect toward establishing the disease and will likely have a larger effect on the progression of the disease in an individual.

Understanding the effect of unintegrated virus on initial viral dynamics can have important clinical significance. For example, productive, unintegrated virus can contribute towards the failure of recent clinical trials of the integrase inhibitor Raltegravir [63, 64]. The R_0 of HIV *in vivo* has been estimated to be around 4-8 [57]. If unintegrated virus contributes 20% towards R_0, then unintegrated virus alone can potentially establish infection in some individuals, although the infection will proceed at a slower pace initially. The analysis presented here can suggest that prescribing an integrase inhibitor in a prophylactic capacity, as has been suggested for high risk groups, may be ineffective due to productive unintegrated virus.

The results here are dependent on parameters that are precisely known but this analysis serves as a starting point for the exploration of the effect of unintegrated HIV. This analysis also provides a mathematical framework and identifies key parameters that should be estimated to understand the effect of uDNA on establishing HIV-1 infection.

CHAPTER 2: Unintegrated HIV and immune responses

Abstract

Recent experiments has shown that cells infected with unintegrated HIV DNA (uDNA) can successfully produce virus. We explore the effect of these unintegrated viruses on asymptomatic phase viral dynamics with a mathematical model. Our model includes synaptic and free-virus transmission, the multiple infection of target cells, and an immune response. We find that uDNA can both increase or decrease set point viral loads and provirus levels, depending on the strength of viral production by cells infected with just uDNA and the strength of the immune response versus these cells. If the viral production by uDNA-only cells is strong and the immune response versus them is weak, then viral loads and provirus levels are higher than if uDNA was inert. This is due to viral production by previously inert cells. However, if viral production by uDNA-only cells is weak and the immune response versus them is strong, then viral loads and provirus levels are lower than if uDNA was inert, despite viral production from a new population of cells. This is due to a form of apparent competition between cells with provirus and cells with only uDNA.

Introduction

Chapter 1 of this dissertation explored the basic reproductive ratio, which is critical at the initial stage of infection, but unintegrated viruses may have an effect on viral dynamics long after the onset of infection. Previous models have ignored unintegrated

viruses altogether because it is assumed that unintegrated viruses are a replicative dead-end. Integration is a complex, multi-stage process and a majority of viruses that enter a cell do not undergo all the steps required to integrate [65]. The relatively large population of cells with uDNA can potentially change infection dynamics. In this chapter, we explore the effect of uDNA on steady states and in the context of an immune response.

In order to study the effect of uDNA we also consider multiple infection of cells and different modes of transmission. Multiple infection is an important factor because it is likely that, in the presence of provirus, episomal uDNA can be more efficiently packaged into new virions. The proteins expressed from the integrated provirus can be packaged along with a transcript from an episomal uDNA in a new virion. Since synaptic and free virus transmission affects the multiplicity of infection, we also include mode of transmission in our analysis.

We employ an ordinary differential equation model similar to the model used in Chapter 1. The description and assumptions of the model apply here. This includes the assumptions in our treatment of synaptic transmission, free-virus transmission, and co-infection. The parameter definitions and estimates are also the same and the analytical results from Chapter 1 still apply to the model that will be used in this chapter.

T-cell effectors, or cytotoxic T-lymphocytes (CTL), are a major component of the immune response and is particularly relevant to HIV because they are the targets of infection. At the start of infection, antigen presenting cells consume virus at the site of

infection, travel to the lymph nodes, and present them to CD4 T helper cells. The helper cells activates the antigen presenting cells, which in turn can activate resting CTL [66, 67]. Originally, there are very few T-cell effectors that are specific to a particular viral epitope but activated CTL increases this population in a process called clonal expansion. Once activated, T-cell effectors can remove infected cells through a variety of mechanisms including cell lysis [68-71].

Immune response dynamics has been studied extensively with mathematical models [46, 72-75]. This includes complex models of antibodies, T-cell responses, immunological memory, and different immune cells [48]. As an initial step, we consider the effect of uDNA in the context of a saturating T-cell effector response. This is an initial linear response that plateaus at some level based on a saturation constant and is more realistic than a simple linear response [75]. In the following model, we represent T-cell effectors as a species that is activated in the presence of virus.

Model

$$\frac{dy_{0,0}}{dt} = l - dy_{0,0} - by_{0,0}v - y_{0,0}\sum_{k=0}^{N}\sum_{l=0}^{M} y_{kl}\sum_{r=0}^{N}\sum_{s=0}^{M} g_{rs}^{kl}$$

$$\frac{dy_{i,j}}{dt} = (1-p)bvy_{i-1,j} + pbvy_{i,j-1} - a_{i,j}y_{i,j} - bvy_{i,j} - w_{ij}^{k}zy_{i,j} + \sum_{k=0}^{N}\sum_{l=0}^{M} y_{kl}\left(\sum_{r=0}^{i}\sum_{s=0}^{j} g_{rs}^{kl}y_{i-r,j-s} - y_{ij}\sum_{r=0}^{N-i}\sum_{s=0}^{M-j} g_{rs}^{kl}\right)$$

$$\frac{dv}{dt} = \sum_{i=0}^{N}\sum_{j=0}^{M} k_{ij}y_{ij} - uv$$

$$\frac{dz}{dt} = \sum_{i=0}^{N}\sum_{j=0}^{M} w_{ij}^{a}z/(z+e)y_{ij} - bz$$

We employ the same model used in chapter 1 of this dissertation but add a species to represent the immune response. The actual immune response is complex and there have been many models built to account for it that is beyond the scope of this dissertation. We will focus on a simple immune response with just a single species. The new species, denoted z, represents CTL effectors. The results presented in chapter 1 of this dissertation still applies to this model because, at the start of infection, the immune system has not mounted a response and the population of CTL effectors is effectively zero. The CTL effectors are activated at rate w^{a}_{ij} and kills infected cells at rate w^{k}_{ij}. In this work, w^{a}_{ij} and w^{k}_{ij} are identical for all cells with more than one provirus and identical for all cells with just unintegrated viruses. The death rate of infected cells (a_{ij}) now represents removal by mechanism other than CTL effectors. We use a saturating activation function to better model immune response and to reduce unrealistic oscillations. Our choice of the parameters w determines the response the immune system to uDNA-infected cells.

For our simulations, the parameter values were chosen based on literature values as cited above. If the values are unavailable, they were varied in our numerical simulations. For figures, the exact parameter values used are stated in the captions. The computational costs rises as N^3 due to the higher number of species and cross interactions between each species as N increases. Due to high computational costs, the maximum number of viruses per cell (N and M) was reduced in simulations after it was determined that lower values of N and M produced qualitatively similar results.

Parameters

Parameter	Value	Units	Notes	Source
N (max unint)	10	none	Maximum unintegrated per cell	
M (max int)	10	none	Maximum integrated per cell	[29]
λ	[10-1E5]	Cells $ml^{-1}day^{-1}$	Uninfected production rate	[76]
d	0.01	day^{-1}	Uninfected death rate	[76]
θ	[3-15]	none	Viral production advantage of provirus	[43]
a_i	0.7, 0.1	day^{-1}	Provirus infected death rate without / with immune response	[11,76]
a_u	0.7/ θ	day^{-1}	Unintegrated infected death rate	
k_i	[10-100]	day^{-1}	Prod. rate of free virus from provirus infected cells	[11,76]
k_u	k_i/θ	day^{-1}	Prod. rate of free virus from unintegrated infected cells	
u	3	day^{-1}	Clearance rate of free virus	[11]
γ	[1E-3 1E-7]	ml $Cells^{-1}$ day^{-1}	Base infectivity for synaptic transmission	
β	[1E-3 1E-7]	ml $Cells^{-1}$ day^{-1}	Infectivity for free virus transmission	[76]
h	[1 25]	none	Average cells successfully transferred per synapse	
δ	0.3	day^{-1}	Chance of competence of unintegrated viruses	UD
p	0.1	none	Chance of integration	UD
c	[1-5]	none	Number of unintegrated viruses required for productivity	[61]
w^{ki}	[1E-4 1E-6]	ml $Cells^{-1}$ day^{-1}	Clearance rate by immune system of cells with provirus	[76]
w^{ku}	w^{ku}/θ	ml $Cells^{-1}$ day^{-1}	Clearance rate by immune system of cells with uDNA only	
w^{ai}	[1E-4 1E-6]	ml $Cells^{-1}$ day^{-1}	Activation rate of immune system by cells with provirus	[76]
w^{au}	w^{ai}/θ	ml $Cells^{-1}$ day^{-1}	Activation rate by immune system of cells with uDNA only	
ϵ	100	Cells ml^{-1}	Saturation constant for immune response	
Initial Uninfected	1E6	None	Initial uninfected	
Initial infected	100	none	Initial singly infected cells (with integrated virus)	
Initial virus	100	none	Initial free virus	

Table 2.1 : Parameters used in steady state simulations. If no source was found then the parameter was varied. UD: from unpublished data.

Initial conditions for the number of initial infected and initial free virus vary widely, based on the mechanism and source of initial infection. For example, a blood transfusion from an infected patient would have initial values that are orders of magnitude higher than from sexual transmission. Initial values were chosen based on a moderate number of initial viruses but numerical simulations show that steady state values were unaffected by varying the initial conditions.

We also set a number of rules for the parameters described in Table 1. We assume that activation is more difficult to trigger than immune-mediated clearance. That is, the number of cells required to activate an equivalent CTL response is always greater or equal to the number required to be targeted by the that CTL response ($\eta^k_{ij} \geq \eta^a_{ij}$). This is generally true since, after the immune response is initially mounted and activated, due to mechanisms such as affinity maturation and clonal expansion, the immune response becomes more efficient. We assume that the production rate of free virus by uDNA-only cells is always less than the production rate of cells infected with a provirus by a factor of θ, unless otherwise specified. This assumption is based on experimental results on the reduced strength of infection from integrase negative HIV [43]. We also assume that the immune clearance rates, death rates, and transmission rates are independent of the number of viruses in the cell (as long as they reach a threshold for production) and are only dependent on the type of virus in the cells.

To ascertain the effect of uDNA in the context of the immune response, we vary the strength of uDNA-only transmission and the immune response versus cells with just uDNA. In the base case, cells with only uDNA do not produce virus and do not activate the immune response. This is case where uDNA-only cells are completely inert until it is infected by a provirus and is standard in other mathematical treatments of HIV. Previous work generally do not explicitly consider uDNA and combines all infected cells into a homogenous population with a single death rate and virus production rate. From the base case, we vary the strength of the immune response relative and viral production rate.

At one extreme, uDNA-only cells produce virus and invokes an immune response that is proportional to the amount of virus produced. For example, uDNA-only cells that produce $1/5^{th}$ of the virus as cells with iDNA will invoke an immune response that is $1/5^{th}$ as strong. This is reasonable if the strength of the immune response depends on how productive the infected cell is.

On the other extreme, uDNA-only cells produce few virus and activates disproportionately strong immune response. In this case, uDNA is marginally productive but still activates an immune response that is as strong as cells infected with a provirus. For example, uDNA-only cells produce $1/5^{th}$ of the virus as cells with iDNA will still invoke an immune response that is equally as strong. Since the production of viable virions is significantly more difficult than initial infection, it is reasonable to assume that a non-productive cell with uDNA can still trigger an immune response as viral epitopes are displayed on the infected cell. Not enough is known about the strength of transmission by uDNA-only cells or the immune response versus them. We explore the effect varying these parameters on the dynamics of HIV.

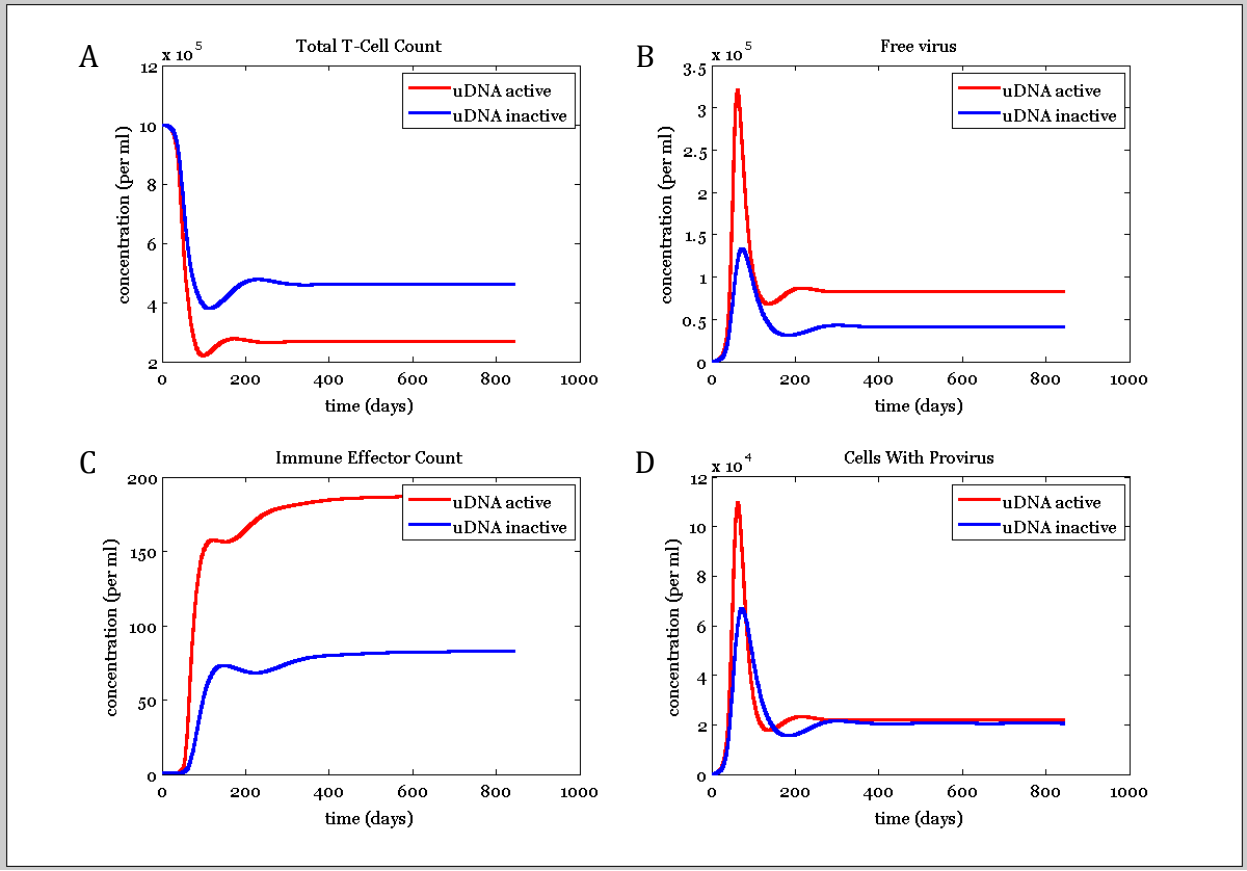


Figure 2.1: Typical simulations at steady state with immune response to uDNA proportional to the amount of virus produced by uDNA-only cells. The ratio of the production disadvantage and clearance rate advantage is 1. Solid lines have productive uDNA, dashed lines are unproductive uDNA – uDNA infected cells do not produce virus, does not die at an increased rate, and is not targeted by the immune system. A) The T-cell count is lower due to increased death rate of active cells with uDNA only. B and C) The uDNA active simulation has a higher free virus load and immune effector count because there is more cells that produce virus and activate the immune system, respectively. D) The number of cells with provirus remains constant despite increased free virus, suggesting most of the new virus production from uDNA-only cells. Parameters are as follows: c=2, β = 1E-6, γ = 2E-7, λ = 10000, u=5 , θ=4 a_i=0.03, a_u=0.03/θ, d=0.004, k_i=10,k_u= k_i/θ , b=0.01, w_{ai}=5E-5, w_{au}= w_{ai}/θ , w_{ki}=5E-5, w_{ku}= w_{ki}/ θ, p=.1, δ=.5, h_{avg}=2, $y_{00}(0)$ = 1E6, $y_{10}(0)$ = 100, $y(0)$ = 10, $z(0)$ = 1

Contribution of uDNA towards set-point viral load

To assess the relative contribution of uDNA to the total viral load, we compare the total number of viruses in the system with productive uDNA versus one without productive uDNA. We vary the immune response and viral production rate of cells with just unintegrated viruses. We find that, if cells with just uDNA produce viruses at a reduced rate and are cleared at an equivalently reduced rate, the total viral load increases as compared to the case with inert uDNA (Fig 2.1 top right). This is simply because there is a population of cells that produce viruses that did not produce virus in the base case. Increasing the strength of uDNA transmission also increases the number of viruses in the system, and therefore increases the number cells infected by proviruses. However, the increase in free virus concentrations cannot be fully explained by an increase in provirus since cells with provirus are only changed by a slight amount (figure 2.1D). There are parameter regions where there is a decreased provirus level but the total free virus concentration increases due to production from uDNA-only cells, which indicates that the increased set-point virus levels is due to production from uDNA-only cells.

When we compare scenarios with active and inactive uDNA, it should be of no surprise that the presence of active uDNA-infected cells could increase set point viral load because more cells are available to produce virus, both through synaptic and cell-free transmission. The magnitude of increase of the viral load depends on the probability of integration and multiplicity of infection. As these factors increase, there are fewer cells

infected with only uDNA and the contribution to viral load at steady state of uDNA-only cells decreases.

As the rate of uDNA-only cell immune clearance increases relative to the strength of virus production by these cells, virus loads decreases. In the extreme case, the virus loads become lower than the case with inert uDNA. (Figure 2.3B) In this case, the immune response is augmented to a degree that, despite increased viral production, the total viral load is lower.

Effect of uDNA on cells with provirus

With a proportional immune response, the number of T cells at equilibrium is lower when uDNA is active versus non-active due to immune-induced clearance of cells with only uDNA (figure 2.1A). When uDNA is not active, the cells that are infected with just unintegrated viruses do not incur an immune response, is not targeted by the immune system, and live longer as a result. When uDNA-only cells are active, the effective infectivity of the entire system is higher and the effective death rate of cells is higher. The death rate of T-cells is higher, on average, and therefore the total T-cell count is lower.

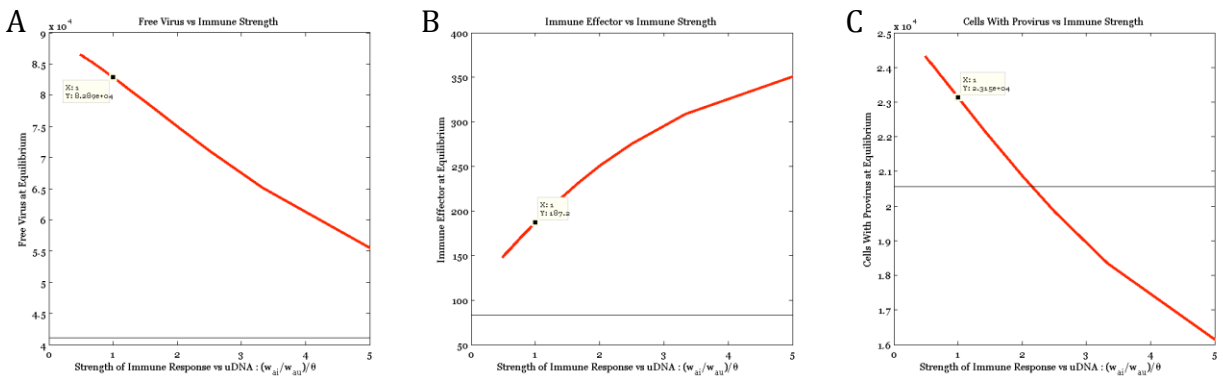


Figure 2.2: Adjusting the immune strength versus cells infected with just uDNA when uDNA is active. Immune strength is measured by immune response (both activation and killing by CTL effectors) relative to virus production. Black lines represent concentration with uDNA inactive. Points marked are proportional immune responses towards cells with uDNA only. A) As the immune response to uDNA becomes stronger, less free virus is produced, B) there are more immune effectors, and C) Provirus is reduced due to apparent competition. Parameters are as follows: c=2, β = 1E-6, γ = 2E-7, λ = 10000, u=5 , θ=5 a_i=0.1, a_u=0.1/θ, d=0.01, k_i=10,k_u= k_i/θ , b=0.01, w_{ai}=5E-5, w_{au}= varies, w_{ki}=5E-5, w_{ku}= varies, p=.1, δ=.5, h_{avg}=2, $y_{00}(0)$ = 1E6, $y_{10}(0)$ = 100, v(0) = 10, z(0) = 1

In addition to changes in free virus and total T-cell counts, when uDNA is active, the cells with only uDNA engage in apparent competition with cells that have integrated viruses. In figure 2.1, the number of cells with provirus remains relatively constant compared to the case with inactive uDNA despite increased free virus. As the strength of the immune response against uDNA-only cells increases, as measured by activation and killing rates of infected cells, the number of cells infected with provirus decreases, supporting the idea that apparent competition is occurring (Figure 2.2C). As a consequence of this apparent competition, as the strength of the immune response against uDNA-only cells increases, the total provirus levels decreases.

In this system, there are some confounding factors that can mask apparent competition in the presence of active cells with uDNA only. With active uDNA, we see an

increased free virus count (Figure 2.1B) and intensity of cell-to-cell transmission, which will tend to increase the number of cells with provirus. However, despite these factors, the presence of active uDNA can be detrimental to cells with provirus, lowering steady state concentrations below what they would be with only active provirus. In our model, this can only be explained by the increased immune response and apparent competition.

Apparent competition has been studied extensively with two or more prey species and a predator [85]. However, in this case, the two types of cells benefit each other when production by uDNA-only cells is active. Without apparent competition, in the absence of an immune response, the presence of active uDNA-only cells increases provirus levels because uDNA-only cells can release and transmit viruses that become proviruses. The reverse is also true – the presence of cells with provirus increases uDNA-only cells. Apparent competition only offsets the increased production if the immune response versus the uDNA-only cells is relatively strong. This would be the case if the total number of viruses produced by uDNA-only cells throughout their lifetimes were significantly less than those of cells infected with provirus. For example, this would be the case if cells with uDNA-only produce viruses at a tenth the rate of iDNA cells, but activates the immune response at a rate that is a half of the rate of iDNA cells. This is also plausible because the requirements for an immune response versus a particular cell are far less than the requirements for active production. A cell infected with uDNA-only can be unproductive but still elicit an immune response due to viral epitopes displayed on the infected cell.

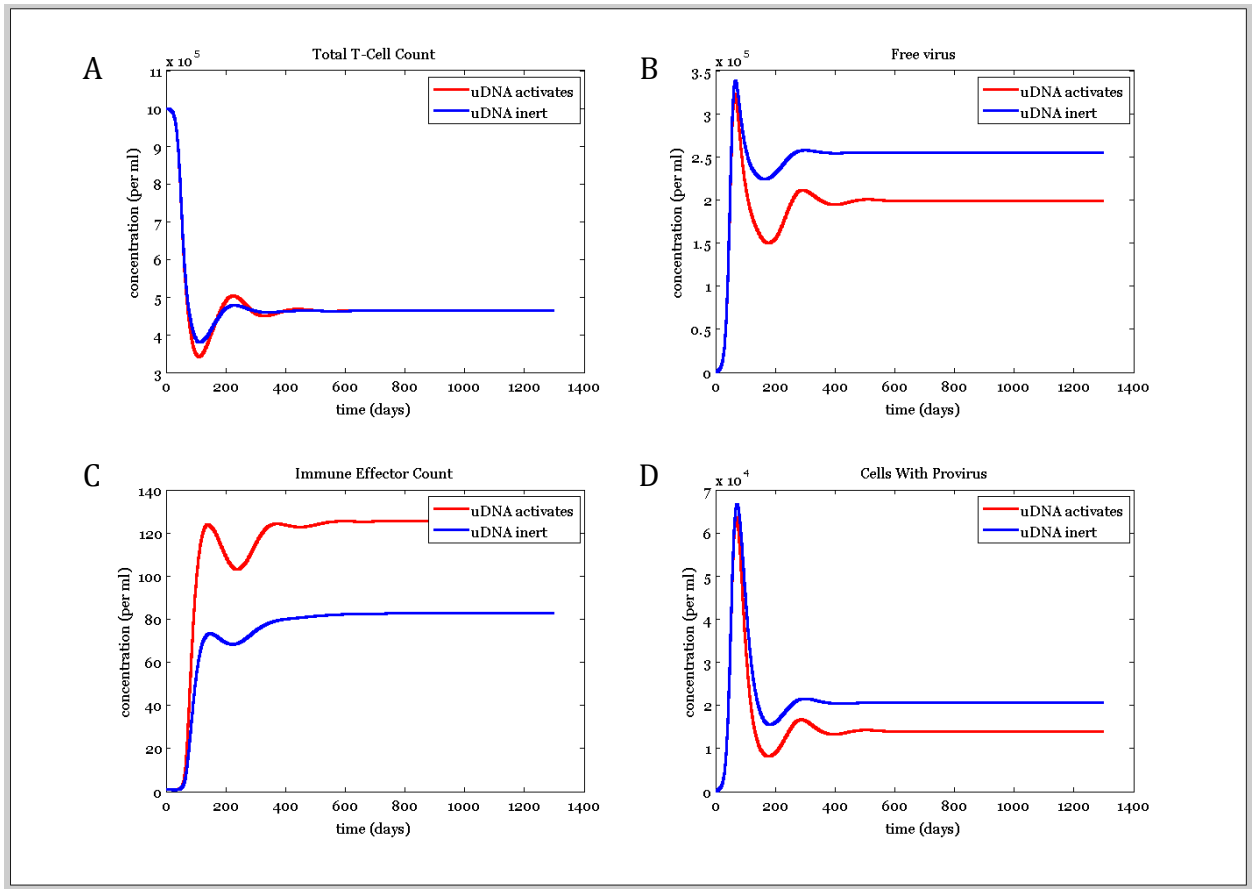


Figure 2.3: Typical simulations at steady state comparing a system where uDNA-only cells cannot produce virus and is ignored by the immune system (inert) versus a system where uDNA-only cannot produce virus but still activates the immune system. A) T-cell counts are equal because there are more uninfected cells in the activated scenario B and C) If uDNA activates the immune system while producing no virus, the free virus levels decrease and the immune effectors increase D) Cells with provirus decrease due to apparent competition from uDNA-only cells. Parameters are as follows: c=2, β = 1E-6, γ = 2E-7, γ(udna) = 0, λ = 10000, u=5 , θ=5 a_i=0.1, a_u=0.1/θ, d=0.01, k_i=10,k_u= 0, b=0.01, w_{ai}=5E-5, w_{au}= w_{ai}/θ or 0 , w_{ki}=5E-5, w_{ku}= w_{ki}/ θ or 0, p=.1, δ=.5, h_{avg}=2, $y_{00}(0)$ = 1E6, $y_{10}(0)$ = 100, v(0) = 10, z(0) = 1

The effect of uDNA as a competitor to iDNA is most clearly seen when uDNA is nonproductive but incurs an immune response (figure 2.3). This is the most extreme case, but the immune response and the number of viruses produced by cells infected with uDNA

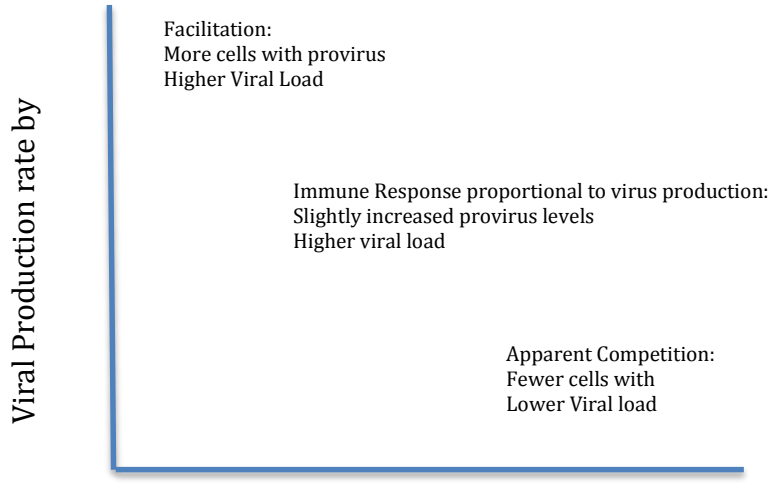


Figure 2.4: Summary of the effect of uDNA.

are not known. It is possible that a cell can be unproductive but, due to the presence of viruses, incur an immune response due to the display of viral antigens on the surface of the cell. In this case, cells with uDNA-only produce a negligible amount of viruses but incur an immune response. Compared to the inert case, the total number of T-cells remains the same (2.3A) but the proportion of productively infected cells decreases (2.3D). uDNA-only cells elicit an immune response, leading to an increase in immune effectors (2.3C) that also target cells with iDNA. In this case, cells infected with uDNA-only decreases the viral load (2.3B) since the immune response is elevated due to nonproductive cells.

The overall effect of uDNA can be summarized in Figure 2.4. If the immune response versus uDNA is low, then cells infected with uDNA will increase the number of infected cells and set point viral load. As the immune response versus uDNA increases, apparent competition is more likely and cells with provirus may decline. Conversely, if viral production rate by uDNA is high, there will be a higher viral load and the infection would be stronger than in the case if uDNA were inert.

Discussion and Conclusions

Starting from the system we used in Chapter 1, we add an immune response and we explore the interaction between uDNA and iDNA in the context of predatory-prey dynamics. We find that uDNA can potentially increase or decrease viral loads, depending on the productivity of uDNA-only cells and the strength of the immune response versus them. If cells with uDNA-only are highly productive and the immune response versus it them is relatively weak, then set point virus loads are higher than they would be with inactive uDNA. If cells with uDNA-only are marginally productive and incur a relatively strong immune response, then set point virus loads can be lower than they would be with inactive uDNA. We also find that the presence of uDNA can decrease or increase provirus levels, depending on the strength of the immune response versus uDNA. If cells with uDNA are productive and incur a weak immune response, then the number of cells with provirus will increase. Conversely, if cells with uDNA are marginally productive and incur a strong immune response, then the number of cells with provirus will decrease.

If uDNA is entirely non-productive and yet incur an immune response, it may have an effect similar to vaccination in the host. However, depending on the strength of replication, it may limit their usefulness in that regard. It is possible, even likely, that cells with uDNA incur strong immune response relative to the amount of virus they produce because they are less efficient at producing virus, while still presenting viral antigens that can invoke an immune response. The case of greater than proportional immune response is even stronger if more than one uDNA is required for production. The requirements for

production are higher while cells infected with a single uDNA may still present antigens. As the threshold number of uDNA required for production increases, the uDNA is more likely to act as a competitor to iDNA rather than enhancing infection.

It should be clear that these are preliminary explorations on the effect of uDNA. Too many parameters are unknown and unmeasured to be able to present a precise description of the dynamics involved. In addition, the parts of the immune system are represented by a single species for simplicity and the true interactions are more complex. However, this work serves to show the theoretical effects of uDNA that has not been explored elsewhere. It also serves as a mathematical framework for the identification of parameters that are key in determining the degree of importance of uDNA in the progression of HIV.

CHAPTER 3 : Multiplicity of HIV infection in the peripheral blood versus lymph

Abstract

We introduce a two compartment HIV model of T-cell dynamics to explain the observed discrepancy in the number of multiply infected cells in the peripheral blood and lymphoid system. Most T lymphocytes in the peripheral blood originate from the lymphoid system, but recent experimental results show that there is few multiply infected T lymphocytes in the blood. Our ordinary differential equation model suggests that the absence of strong synaptic transmission in the peripheral blood tends to create many singly infected cells – reducing the relative proportion of multiply infected cells in the blood. Although this simple mechanism can explain much of the difference between the two systems, we also conclude that there must be some additional mechanism that removes or prevents highly multiply infected cells from entering the peripheral blood. We employ an agent-based model and determined that increased death rates of highly multiply infected cells, increased death rates for older T-cells, and spatial differences in release rates can preferentially reduce the number of highly multiply infected cells in the peripheral blood system.

Introduction

HIV targets and infects cells with the CD4 marker including helper T cells, macrophages, and dendritic cells. These cells reside both in the peripheral blood and in the organs of the lymphoid system including the thymus, lymph nodes, and spleen. In this paper, the lymphoid system is considered the lymphoid organs including the lymph nodes and spleen but excludes free-flowing lymph fluid. Many cells in the lymphoid system are infected with multiple copies (average of 3.2 and up to 8) of HIV-1 provirus during infection [29, 77]. Recently, a paper has been released with seemingly contradictory results – most infected CD4+ cells in the blood have only 1 virus per cell [33]. The results are summarized in figure 3.1, taken from Jung 2002 and Josefsson 2011. Although the two studies were performed in cells from different organs, these two locations are connected - the organs of the lymphoid system (including the spleen) releases mature CD4+ T lymphocytes into the blood. This work seeks to determine what may cause the observed disparity between the multiplicity of viruses in cells of the blood and lymphoid system.

HIV in CD4+ T cells can spread in two different ways. The first mode of transmission is free virus or cell-free transmission, where productively infected cells release HIV virions that infect other cells. This is the typical method considered in most mathematical models of HIV. A productively infected cell releases thousands of virions before it is cleared but each infected cell only infects about 8 other cells, on average [57]. Thus, free virus infection is an inefficient process and a majority of viruses released are degraded or cleared prior to producing a productive offspring cells. The second mode of transmission is via a viral

synapse that can form between two T-cells [39, 78]. Instead of a single virus entering targets, a productively infected cell can transmit many viruses directly into a target cell, avoiding the humoral response. In contrast to free virus transmission, most of the viruses that are produced enter a susceptible cell and this may be more efficient than free virus transmission. The exact mechanism of viral synapse formation is unclear, but studies have shown that it takes approximately 30 minutes for a synapse to form and viruses to be transferred between cells [39]. Clearly, the cells must remain in close proximity to each other during this process.

There are distinct morphological differences between the peripheral blood and the lymphoid system that may lead to different infection dynamics [79]. Key differences include the number, density, and mobility of T-cells in the two compartments. Approximately 98% of of CD4+ T lymphocytes and precursors are found in the lymphoid system so we would expect that the multiplicity of infection in the blood would mirror that in the lymphoid system [80-81]. CD4+ cells in the peripheral blood are more mobile and are sparser than in the spleen. Since mobile cells are less likely to engage in efficient synaptic transmission, synaptic transmission would be less efficient in the blood [82]. We use mathematical models to explore the effect of these modes of transmission on HIV dynamics and multiplicity of viruses in infected cells.

Mathematical models of HIV infection have traditionally treated the infected environment as a single, homogenous system. Here, we use a two-compartment ordinary

differential equation model with synaptic and free virus transmission to understand the observed disparity between the multiplicities of HIV proviruses in cells.

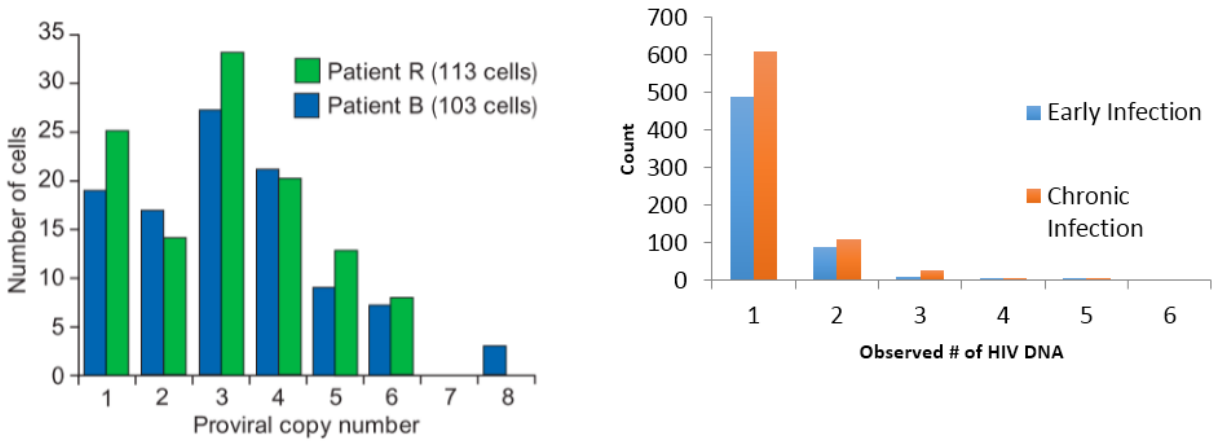


Figure 2.1: Distributions of multiplicity of infectivity in the spleen (left) and blood (blood). The distributions are qualitatively different – a majority of cells in the spleen are multiply infected while a majority of the cells in the blood are singly infected. Data summarized from [15] and [30]

Two Compartment ODE model

In the blood:

$$\frac{dy_0^B}{dt} = I^B - dy_0^B - by_0^B v^B + ry_0^L$$

$$\frac{dy_i^B}{dt} = bv^B y_{i-1}^B - a_i y_i^B - bv^B y_i^B + ry_i^L$$

$$\frac{dy_N^B}{dt} = bv^B y_{N-1}^B - a_N y_N^B + ry_N^L$$

$$\frac{dv^B}{dt} = \sum_{i=0}^{N} k_i y_i^B - uv^B$$

In the lymph:

$$\frac{dy_0^L}{dt} = l^L - dy_0^L - by_0^L v^L - y_0^L \sum_{k=0}^{N} y_k^L \sum_{r=0}^{N} g_r^k - ry_0^L$$

$$\frac{dy_i^L}{dt} = bv^L y_{i-1}^L - a_i y_i^L - bv^L y_i^L - \sum_{k=0}^{N} y_k^L \left(\sum_{r=0}^{i} g_r^k y_{i-r}^L - y_i^L \sum_{r=0}^{N-i} g_r^k \right) - ry_i^L$$

$$\frac{dy_N^L}{dt} = bv^L y_{N-1}^L - a_N y_N^L - bv^L y_N^L - \sum_{k=0}^{N} y_k^L \left(\sum_{r=0}^{N} g_r^k y_{N-r}^L \right) - ry_N^L$$

$$\frac{dv^L}{dt} = \sum_{i=0}^{N} k_i y_i^L - uv^L$$

Parameters:

d : Death rate of uninfected CD4+ T cells, can also be called a_{00}

λ_B : birth rate of uninfected cells in the blood

λ_L : birth rate of uninfected cells in the lymph

β_B : Infectivity – A measure of the effectiveness of free-virus transmission in the blood

β_L: Infectivity - A measure of the effectiveness of free-virus transmission in the lymph

K_i : Virus replication parameter – rate that an infected cell releases free virus

u: Clearance rate of free virus

a_i Clearance rate of infected cells

r: release rate from lymphoid system to peripheral blood

γ^k_l the probability that a cell with k viruses successfully transfers l viruses

The model seeks to explain the following observed phenomenon:

1. A majority of the infected cells in the lymphoid system are multiply infected

2. A majority of the infected cells in the peripheral blood is singly infected
3. There are very few highly multiply infected cells in the blood
4. There are significantly more cells in the lymphoid compartment
5. Parameters must be close to experimentally determined values, if known.
6. Very few maximally infected cells

We model the entire peripheral blood system and the lymphoid system as separate compartments. Although this can be converted to concentrations, it would require renormalizing the parameters in each compartment due to the differences in size of compartments, and therefore, the concentration of the species in each system.

The average number of viruses per cell is an extremely poor measure of the differences between compartments. The observed average number of viruses per cell can br obtained by increasing or decreasing the infectivity (k, β, γ, death rates, λ) in each compartment. However, this often leads to unrealistic numbers of maximally infected cells in the lymphoid system.

Parameter Values

Parameter	Definition	Value (Blood)	Value (Lymph)	Notes	Source
N	Max # of provirus per cell	8	8	1	[15]
u	Free virus removal rate	0.5	0.5	1	[7]
d	Uninfected death rate	0.01	0.01	1	[7]
a	Infected death rate	0.7	0.7	1	[7]
k	Free virus production rate	20	20	1	[37]
λ	Uninfected birth rate	5E8	2E10	2	[36]
r	T-cell release rate	N/A	0.03	1	
γ	Synaptic Infectivity	N/A	1E-10	1	
β	Cell-free Infectivity	1E-11	1E-12	1	
Initial Uninfected		1E10	1E11		[36]
Initial Infected		100	0		[36]

Table 3.1 : Parameters used in two compartment ODE model. Notes: 1: Always equal between compartments. 2: Varied Paramete. Uncited parameters were varied.

As in all models of complex biological systems, we take some simplifying assumptions. We assume that the efficiency of synaptic transmission is independent of the number of viruses in either the infected cell or the target. That is, $\gamma_{s,}$ the probability of successfully transferring s viruses, is constant all cells. This is a simplifying assumption that we take as a first step towards understanding the system. A non-constant infectivity does not change the qualitative results of this work. The parameter γ_s is the probability of transferring viruses that eventually become integrated proviruses. In fact, dozens or hundreds of viruses can be transferred per synapse, but very few become active proviruses [29, 83]. We also assume that the number of viruses transferred successfully per synapse is normally distributed. The distribution of the number of viruses transferred is unknown,

but a non-normal distribution of the number of viruses per cell has not been observed in the lymph.

Hypotheses on cause of differences between compartments

There can be many explanations for the observed disparity between the number of infected cells in the lymph and in the blood. For example, increased immune response in the blood against multiply infected cells, increased free virus infectivity rates for cells that have already been infected, and morphological differences between stationary and mature T-cells that are released from the lymph. These possibilities are unexplored and any or all of them may be in effect during HIV infection. We present another plausible explanation for the observed distributions that we explore through our models. We compare the simulated, steady-state distributions of multiply infected cells in the blood and lymph to the observed data to show that the characteristics present in our model is sufficient to reproduce the data.

Relative frequency of singly versus multiply infected cells

The first observation we wish to address is that a majority of infected cells in the lymph are infected with multiple viruses while a majority of infected cells in the peripheral blood is infected with just one virus. One might expect that both the lymph and blood systems would contain relatively equal ratio of multiply and singly infected cells. Our simulations suggests that, except for highly multiply infected cells, it is possible to produce the observed ratio of multiply infected versus singly infected without invoking any

additional multiplicity-induced mechanisms. For example, increased death rates, increased infectivity, or reduced susceptibility of multiply infected cells is not necessary to reproduce the observed ratio between singly and multiply infected cells (Figure 3.2).

The observed ratio of singly and multiply infected cells in the spleen and the peripheral blood can be explained as follows. Uninfected, singly infected, and multiply infected T-cells are released from the lymph into the blood. Uninfected T-cells are also released directly from the bone marrow, where they are produced, into the blood. In the blood, the dominant mode of transmission is cell-free transmission and, since a majority of the cells are uninfected, most cells that are infected become singly infected. Since these singly infected cells in the blood vastly outnumber the multiply infected cells that originate from the spleen, any sample of peripheral blood T-cells would have a majority of singly infected cells.

The above analysis provides a few testable predictions. If sample sizes were large and the assay sensitive, we would expect to see some multiply infected cells in the blood. In addition, as the immune production falls, we would expect to see a relatively higher proportion of multiply infected cells in the blood since there are fewer uninfected cells for cell-free transmission. We can also use the fitted parameters to estimate some unknown parameters of synaptic transmission.

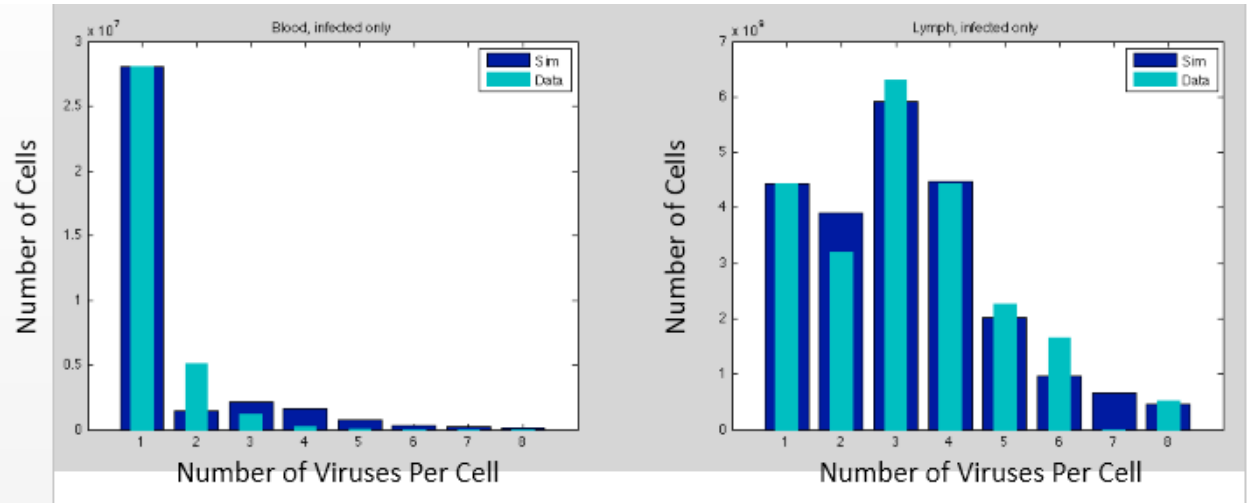


Figure 3.2: Best parameter fit reproduces the distributions observed. Here, r=0.03, u=.5, kb=kl=20, β_L=1.8E-13, β_B=1.3E-13, γ=6E-11, average number of viruses successfully transferred: 2.35, λ_b = 2E8, λl = 1E9.

	Meaning	Value (Blood)	Value (Lymph)
γ	Synaptic Transfer Infectivity		2.8E-11 cells^{-1} day^{-1}
β	Free Virus Infectivity	1.14E-11 cells^{-1} day^{-1}	3.5E-11 cells^{-1} day^{-1}
λ	Uninfected cell production rate	5.7E8 cells^{-1} day^{-1}	2.5E9 cells^{-1} day^{-1}
	Viruses Transferred Per Synapse		2.35
	StdDev of Viruses Transferred		0.78

Table 3.2: Parameter estimates based on least squares fitting Under these parameters, R_0 is approximately 3, which is within the range of published values [37].

Average number of viruses transferred per synapse

Under the assumption that the number of viruses transferred successfully per synapse is normally distributed, we can estimate the average number of viruses

transferred per synapse in the lymphoid system. We fit the parameters of our ODE system to the observed distributions. Based on this method, the average number of viruses successfully transferred per synapse was 2.35 with a standard deviation of 0.78. To ensure that this method is independent of any error in parameter estimation, we varied each of the parameters β, γ, λ_b, and λ_l by an order of magnitude around the optimum value and determined the best-fit value for the average number of viruses transferred. These simulations were all consistent with an average virus transferred of between 2.5 and 3.5, with standard deviation of approximately 0.8. In hindsight, this value can be estimated from the graph in figure 3.1. The average number of viruses per multiply infected in figure 3.1 is 3.2. However, the average of 3.2 includes any multiply infected cells that may have been created by cell-free transmission and the mean number of viruses transferred via the synapse will likely be lower.

Strength of synaptic versus free virus transmission in the lymph

One question of interest is the strength and importance of free virus and synaptic transmission towards viral loads. To measure the strength of transmission we look at 2 different metrics. The first measure is simply the total number of new infections of uninfected cells produced by each mode of transmission since the start of infection. The second measure is the total number of viruses produced by each mode of transmission that successfully enters a cell. For the latter two measures, we pick a time point after equilibrium has been reached and compare the total number of new infections and viruses produced since the start of infection. In our model, both metrics will monotonically

increase over time. However, the ratio of these measures for synaptic versus free-virus transmission remains constant.

Each measure of the strength of transmission yields the same result: synaptic transmission is stronger than cell-free transmission in the lymphoid tissue. The number of new infections suggests that the strength of synaptic transmission in the lymphoid tissue is approximately an order of magnitude stronger than free virus transmission. The number of viruses successfully transferred agrees: synaptic transmission transfers approximately 30x the number of successful viruses as the free virus transmission (See figure 3.3). Since the fitted parameters are based on steady-state data, these rough estimates of the strength of transmission are only valid after the acute phase of HIV infection.

Dependence on parameters

A key question in modeling is the dependence of the results on any parameter values. We varied parameters individually within estimated physiological ranges. If these ranges were unknown, we varied the parameter by approximately an order of magnitude in either direction. We then fit our model to the data to estimate the unknown parameters. Regardless of the parameters we chose, the relative strength of transmission and average number of viruses transferred remained consistent (See figure 3.4). Based on the first measure, synaptic transmission was between 7 and 15x stronger than free-virus transmission at equilibrium in the lymph. The average number of viruses transferred per synapse was between 2.3 and 2.7 in all parameter ranges. Thus, our results are relatively robust against variation in parameter values.

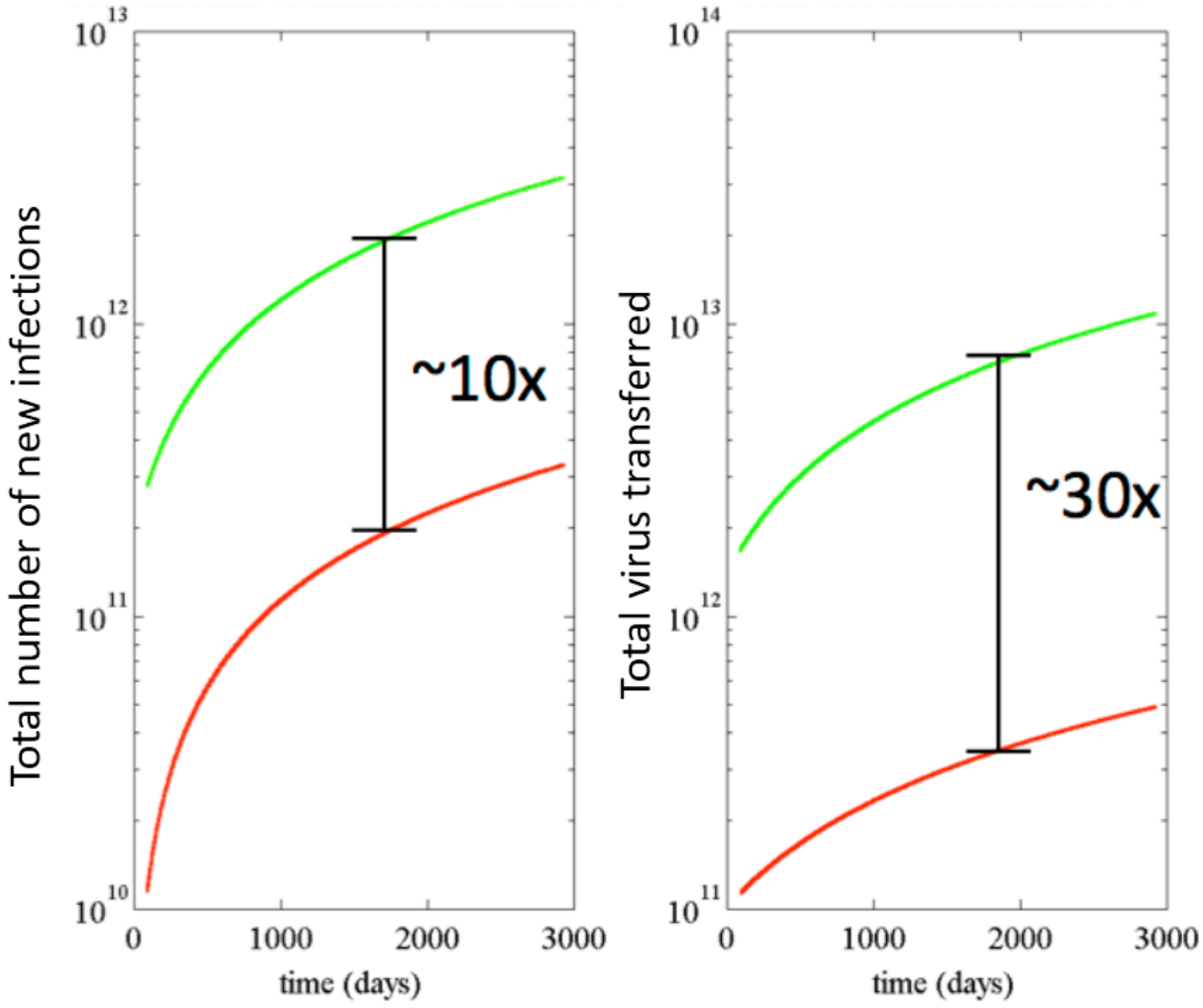


Figure 3.3: Relative strength of each mode of transmission. (left) Metric is the total number of new infections (right) Metric is the total number of viruses successfully transferred. Based on these metrics, the strength of synaptic transmission is approximately an order of magnitude stronger than cell-free transmission in the lymph at equilibrium.

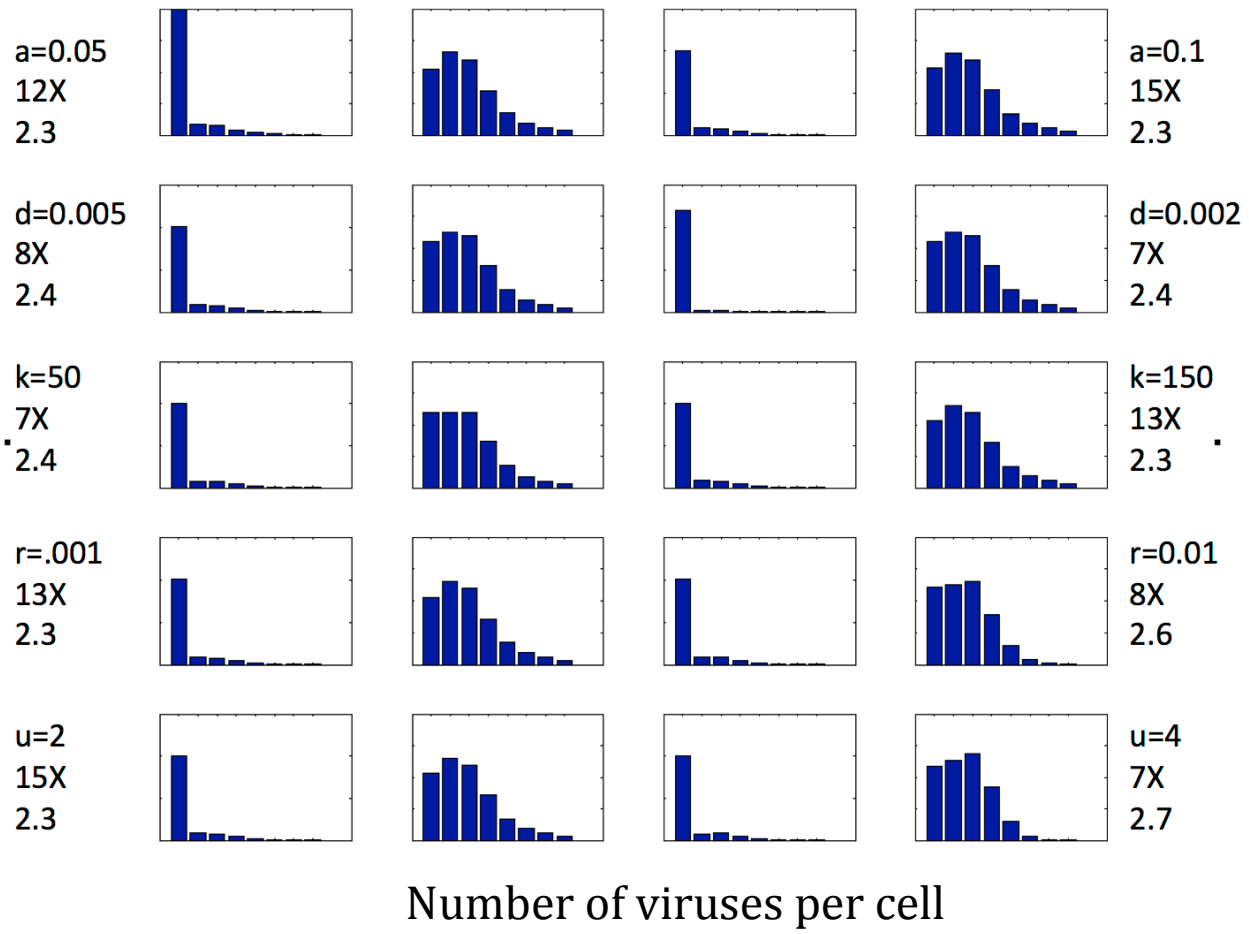


Figure 3.4: Changing parameters does not change the average number of viruses transferred per synapse or the relative strength of each mode of transmission significantly. The parameters were varied based on physiological ranges, if known. Otherwise, the parameters were varied by approximately an order of magnitude.

Distribution of highly multiply infected cells

Despite explaining the ratio between singly and multiply infected cells in both systems, it is not possible to produce the observed distribution for highly multiply infected cells without invoking any additional mechanisms. In figure 3.2, the distribution of

multiply infected cells in the lymph is not statistically different from experimental values (Kolmogorov Smirnov test, p-value = 0.63). However, the distribution of multiply infected cells in the blood is statistically different from experimental values (KS test, p-value < 0.0025). This is partially because the sample size for those experiments was large and any mathematical model cannot account for all phenomena in a biological system. The difference between the simulated and experimental distributions in the blood is due to the presence of highly multiply infected cells. If we remove these highly multiply infected cells from the blood, the distribution of multiply infected cells in the blood from our model is not significantly different from experimentally observed distributions (KS test, p=0.15). Investigators were not able to find any T-cells in the blood with more than 3 viruses in a sample of over 6000 T-cells [33]. In our model, if we assume that the lymph is at equilibrium, it can be shown that show that there will always be highly multiply infected cells in the blood if there is a source of these cells from the lymph. Thus, there must be some other mechanism where highly multiply infected cells are either removed from or prevented from entering the peripheral blood system.

There are a few plausible mechanisms that may contribute to the lack of highly multiply infected cells in the blood. These cells may be cleared at a higher rate, both through immune and apoptotic mechanisms. In addition, highly multiply infected cells may not survive the release from the lymphoid system. There may be spatial differences in release rates of T-cells from the spleen that favors the release of singly infected cells. We explore these possibilities in a stochastic model that will be introduced later.

Reduced / Steady State Model

A constant concern with mathematical models and data is over-fitting. We explored a simplified model with fewer parameters to address the issue. In the following model, we ignore the initial stages of infection and only consider the steady state, which is most relevant to the experimental distributions we observe. The new model also does not explicitly track free virions, further reducing the number of parameters in our model.

In the blood:

$$\frac{dy_0^B}{dt} = a\sum_{i=0}^{N} y_i^B - b^B y_0^B\left(\sum_{i=1}^{N} y_i^B\right) + r^L y_0^L - r^B y_0^B$$

$$\frac{dy_i^B}{dt} = b^B\left(\sum_{i=1}^{N} y_i^B\right) y_{i-1}^B - a y_i^B - b\left(\sum_{i=1}^{N} y_i^B\right) y_i^B + r y_i^L - r^B y_i^B$$

$$\frac{dy_N^B}{dt} = b\left(\sum_{i=1}^{N} y_i^B\right) y_{N-1}^B - a_N y_N^B + r y_N^L - r^B y_N^B$$

In the lymph:

$$\frac{dy_0^L}{dt} = a\sum_{i}^{N} y_i^L - b y_0^L\left(\sum_{i=1}^{N} y_i^L\right) - y_0^L \sum_{k=0}^{N} y_k^L \sum_{r=0}^{N} g_r^k - r y_0^L + r^B y_0^B$$

$$\frac{dy_i^L}{dt} = b^L\left(\sum_{i=1}^{N} y_i^L\right) y_{i-1}^L - a y_i^L - b^L\left(\sum_{i=1}^{N} y_i^L\right) y_i^L - \sum_{k=0}^{N} y_k^L \left(\sum_{r=0}^{i} g_r^k y_{i-r}^L - y_i^L \sum_{r=0}^{N-i} g_r^k\right) - r y_i^L + r^B y_i^B$$

$$\frac{dy_N^L}{dt} = b^L\left(\sum_{i=1}^{N} y_i^L\right) y_{N-1}^L - a y_N^L - b^L\left(\sum_{i=1}^{N} y_i^L\right) y_N^L - \sum_{k=0}^{N} y_k^L \left(\sum_{r=0}^{N} g_r^k y_{N-r}^L\right) - r y_N^L + r^B y_N^B$$

Parameter Values:

Parameter	Definition	Value (Blood)	Value (Lymph)	Source
N	Max # of provirus per cell	8	8	[15]
a	Infected death rate	0.5	0.5	[7]
r^B	Blood to lymph release	N/A	Varied	
r^L	Lymph to blood release	N/A	$r^B/100$	
γ	Synaptic Infectivity	N/A	Varied	
β^B	Cell-free Infectivity	1E-11	Varied	
β^L	Cell-free Infectivity	1E-11	Varied	

Table 3.2 : Parameters used in two-compartment steady state ODE model. Uncited parameters were varied.

We assume that the total number of cells in the system remains constant. Infected and uninfected cells that die are replaced at a commensurate rate. This is approximately the case during the chronic phase of HIV infection. Cell-to-cell infectivity is assumed to follow a normal distribution and the infectivity values for each number of viruses transferred is determined by the mean and standard deviation.

The parameter r^B was varied randomly and the optimal parameters were determined via a steepest descent algorithm. The release rate was fixed in each set of simulations because an optimal fit would always tend to minimize the release rates. For each r^B, we performed a series of simulations with random initial parameter estimates and random initial conditions for the number of uninfected and infected in each compartment. The best fit to the data was determined by a least squared criterion alone. A minority of the simulations produced negative optimal parameter estimates and was discarded. Simulations are coded in Matlab and C++.

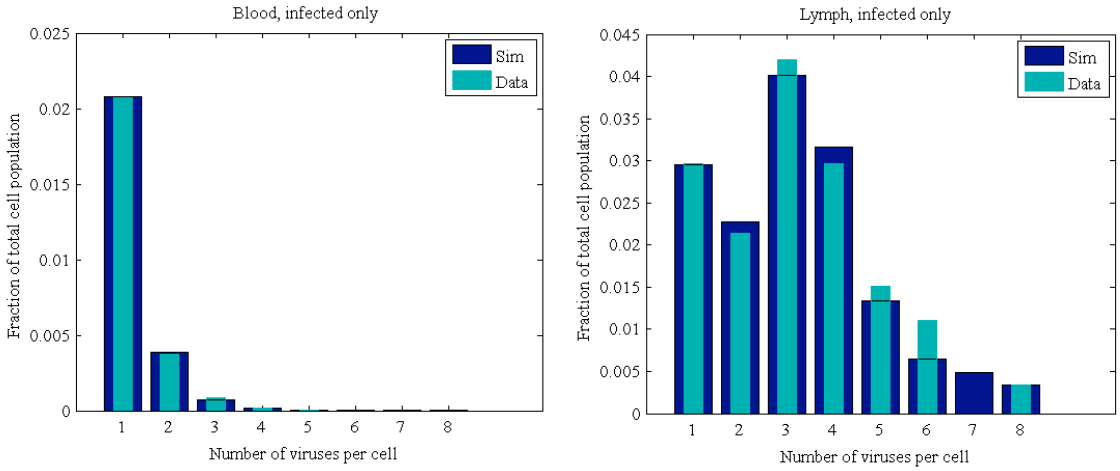


Figure 3.5: Sample optimal fits to data. Bimodal distribution is preserved. Units on y-axis are as a fraction of cells in the entire system.

Parameter Estimates

The exact optimal values for free virus infectivity in the blood can vary as the values for the release rates vary. However, there are some trends that are consistent throughout all the simulations.

Parameter	Average	Min	Max	StdDev
Free-Virus Infectivity in Blood	30.6	2.04	94	33.92
Free-Virus Infectivity in Lymph	0.12	0.05	0.17	0.03
Cell-to-cell infectivity	0.40	0.33	0.52	0.05
Mean viruses transferred per synapse	2.37	2.34	2.47	0.04
Standard deviation of virus transferred	0.83	0.77	1.03	0.09
Ratio of Viruses transferred	8.38	6.13	15.11	2.49

Table 3.3: Parameter estimates from steady state model. Parameter estimates are from 813 simulations with varying initial conditions.

In all the simulations, the number of viruses transferred via cell-to-cell transmission in the lymph is about an order of magnitude higher than number of viruses transferred via free virus transmission. This suggests that, at steady state, the strength of cell-to-cell transmission is stronger than free virus transmission. This is consistent with the more complicated model explored previously. If free virus transmission tends to create singly infected cells then the distribution of multiply infected cells in the lymph suggests that a majority of infections are via cell-to-cell transmission.

The estimates of the parameter representing free virus infectivity in the blood are high and can be misleading. Our steady state simplification of the above model does not account for free virus explicitly. We use the simplification that the number of infected cells in a compartment is a reasonable proxy for the free virus levels. However, since there is an exchange of virions between the lymph and blood, the true virus levels in the blood is higher than the number of infected cells would suggest. Since the lymphoid system contains approximately 50 times the CD4+ T-Cell population then in the peripheral blood, the cells in the lymph likely produce many of the virions in the blood. Since our model does not account for this, the estimates for free virus infectivity in the blood is likely high since it assumes all the new infections in the blood are due to other cells in the blood, rather then from the free virus flowing from the lymph. We believe the estimates for the strength of infection in the lymph is more accurate because transfer of free virus from the blood to the lymph is less likely since the population of infected cells in the lymph is much higher. New free-virus infections in the lymph are likely caused by virions produced by infected cells in the lymph, rather than from virions produced in the blood. In addition, infectivity

parameters may seem high because they are dependent on cellular density and volume, and can't be compared directly to infectivity parameters in different environments.

The mean number of viruses transferred is approximately 2.37 while the standard deviation is 0.83. This is consistent with the more complicated model and is evident from the distribution of multiply infected cells observed.

Effect of Release Rates

r^L	r^B	r^B / r^L	β^B	β^L	γ	γ / β^L
5E-4	5E-3	10	6.74	0.13	0.40	**7.22**
5E-4	0.01	20	12.8	0.12	0.39	**7.28**
5E-4	0.05	100	62.76	0.10	0.36	**8.74**
5E-3	0.05	10	6.82	0.10	0.39	**9.30**
1E-3	0.01	10	6.71	0.13	0.40	**7.27**

Table 3.4: Parameter estimates from varying release rates. The relative strength of synaptic and free virus transmission remains approximately the same regardless of release rates.

Regardless of release rate, the model fits the data fairly well. The multiplicity of infection in the lymph and blood, the relative size of the compartments, and the bimodal distribution in the lymph are replicated with the simulations. Although the parameters estimates do vary, some attributes are consistent throughout all the simulations. Release rates control the number of cells in each compartment at steady state. Since infectivity is a complex parameter, it is affected by many factors including the number of viruses and cells

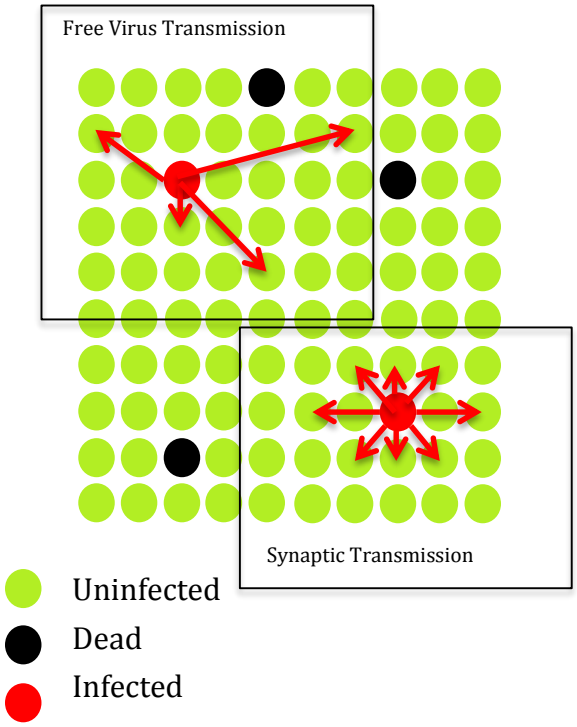
in each compartment and varies with the ratio of release rates. However, the relative strength of each mode of transmission in the lymph is consistent regardless of the both the relative and absolute magnitude of the release rates.

The explanation for the difference in multiplicity of infection in the blood and lymph remains the same as in the more complex model. Multiply infected cells are released from the lymph but free virus infection creates more singly infected cells in the blood.

Cellular Automaton Model

To explore possible mechanisms for the absence of highly infected cells in the blood, we use a cellular automaton model to explore the effects of increased death rates, cellular longevity, and space. With this model, we can explore any spatial effects and any effects due to the age of the cells. The cells were modeled in two different compartments and, at the onset, small portions of the cells are seeded with viruses. For each time step and at each grid point, there is a probability that a cell infects another cell through cell-free transmission, a probability of death, and, in the case of the lymph compartment, a probability of release into the blood and synaptic infection. Synaptic infection is modeled as a release of viruses into cells that are neighbors to an infected cell. Since the lymph is a dynamic system, we allowed synaptic transmission between cells in a certain neighborhood, not only to nearest neighbors. If a grid point is empty, there is a probability that a new cell is born into that grid point. Since this system is stochastic, the probability of each event was estimated by fitting the average of the steady states of multiple simulations

to the experimental distributions. With this system, we can explore the effects of space and cellular longevity.



1. Randomly seed both compartments with infected cells

2. Each time step:
Randomly select coordinates, selected coordinates have a probability of:
-Birth
-Death
-Free Virus Infection
Lymph only:
-Synaptic Infection (within a neighborhood)
-Release to blood

Figure 3.6: Diagram of cellular automaton model. The probabilities of each event at each time step were roughly fitted to the experimental distributions. Synaptic transmission was modeled as the movement of viruses between an infected source cell within a defined distance.

Higher death rates, lower release rates, and effect of cellular longevity

Starting from the base system, we introduce higher death rates for multiply infected cells. Although this assumption is not necessary to reproduce distributions similar to the ones in both compartments, it may be the case in reality. The average death rate of all infected cells, which has been measured extensively, remains approximately the same but multiply infected cells have a higher death rate than singly infected cells. Unsurprisingly, this selects against highly multiply infected cells and produces a distribution of multiply infected cells that is closer to the observed distributions. If we assume that death rates

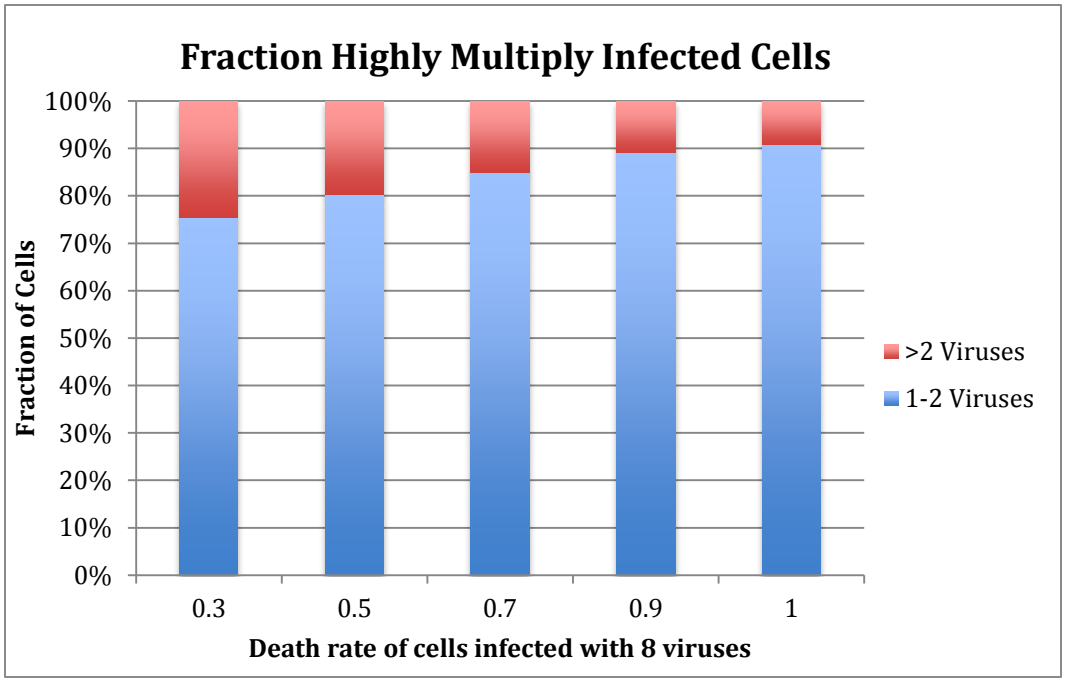


Figure 3.7 Increased probability of death of multiply infected cells decreases their proportion in the blood. Here, death rates vary linearly with the number of viruses in the cell. Each bar is the average of 100 simulations

increase with the number of viruses per cell, then as we increase the death rates of highly multiply infected cells, we see less of these cells in the blood.

The journey from the lymph the blood may select against cells that are already compromised by multiple infections. Not surprisingly, if we preferentially lower the release rates of highly multiply infected cells then we see less highly multiply infected cells in the blood. The effect of this is similar to the effect of increasing death rates.

If we increase the death rates for older cells, we also select against highly multiply infected cells in the blood. Although older cells are rarer in our base simulation because cell death is a Poisson process, we add an additional increase in death rate for multiply infected cells. Older cells tend to be more highly multiply infected because they have survived long enough to undergo more than one infection event. If older cells die at a faster rate, then

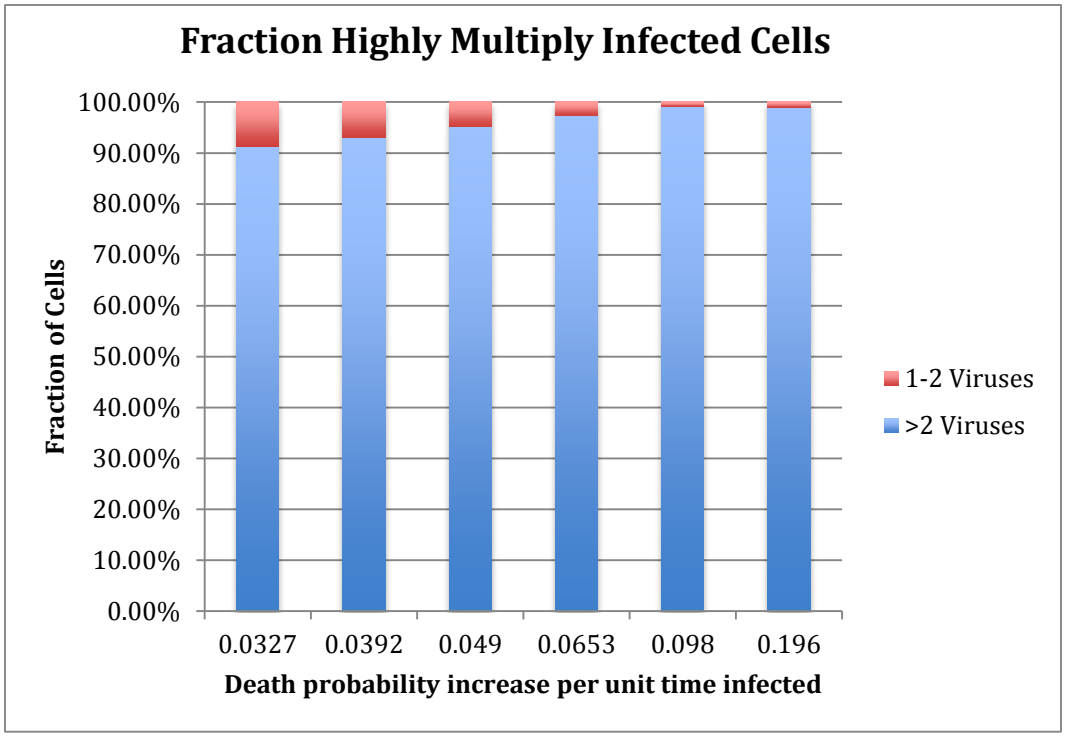


Figure 3.8: Increased death probability of older cells decreases the proportion of highly multiply infected cells in the blood. Here, the probability of death for any single cell is a linear function of the age of the cell. Since older cells tend to be more multiply infected, this selects against highly multiply infected cells. Each bar is the average of 100 simulations.

this could help contribute to why we do not observe highly multiply infected cells in the blood.

Spatial Effects

Finally, we explore spatial effects that reduce the number of highly multiply infected cells in the blood. Although the cells in the lymph move dynamically, it is not a perfectly mixed system and there is some spatial structure. For example, cells near efferent blood vessels may have a higher chance to be released than the ones that are further away. These cells may be less

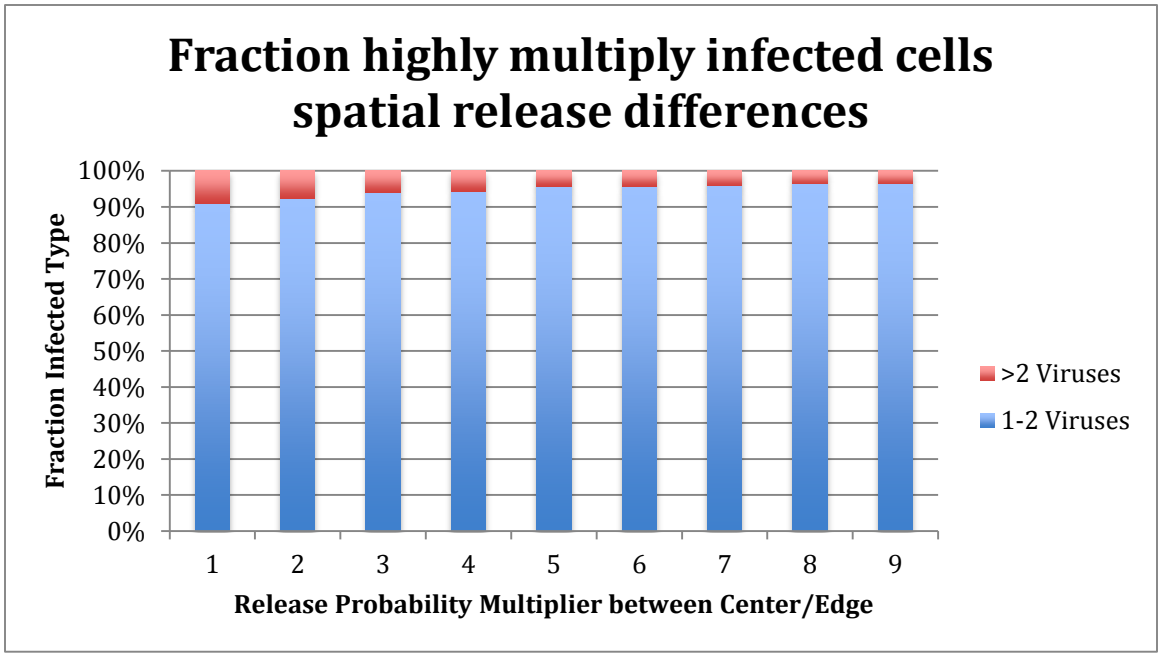


Figure 3.9: Spatial effects – lower release rates in the center. Differences in release rate due to space may also decrease the number of highly multiply infected cells in the blood. Here, the release probability at the edge and center of the domain was modeled by an inverse normal distribution. The release rate at the center was lowest and the release rate at the edge was the highest. The magnitude of the difference between the edge and the center was varied. Each bar is the average of 100 simulations.

multiply infected because they are not surrounded by as many infected cells and may be less stationary. To explore this effect, we can vary release rates in the lymphoid compartment. Cells closer to the center of the lymphoid compartment will have a lower release rate than the cells at the edge. The cells closer to the center of the compartment may be more multiply infected because they remain in the lymphoid system for a longer period of time and are surrounded by more infected cells, on average.

Our simulations show that spatial differences can cause the observed differences (figure 3.9). Parameters are the same between multiply and singly infected cells but the release probability at each grid point. By varying release rates, we can reduce the number of multiply infected cells in the blood without increased death rates for multiply infected or older cells.

New Distribution of Singly Infected Cells in the Lymph

Recent experiments on the multiplicity of infection of HIV have found that, in lymph nodes, cells may be singly infected [34]. This is somewhat contradictory with other results in the literature, which suggest that cells in the spleen, which have a similar anatomical structure and function as the lymph nodes, are multiply infected. Our models suggest a potential cause of this disparity is a change in strength of synaptic transmission. This can be due to differences in viral strains or host factors such as T-cell density mobility in the spleen versus the lymph nodes.

Starting from optimal parameters to fit the Jung et. al data in the steady state model, we refit to the data available on singly infected cells to determine the magnitude of change required to switch between the two types of distributions. This was done for a variety of initial conditions and initial parameter guesses. We assume that the mean and standard deviation of the number of viruses transferred have the same values as the optimal fits for the high MOI data. Our models suggest a decrease in synaptic transmission of approximately 2/3rds produces a majority of singly infected cells versus a majority of multiply infected cells in the lymph. Decreasing the number of viruses transferred per synapse has the same effect as reducing the strength of synaptic transmission.

Starting from optimal fits for the high MOI data, our model requires a drastic change (approximately a 30 times decrease) in synaptic transfer strength to reach the optimal fits for the low MOI data. However, as stated above, an approximate fit only requires a

relatively small decrease in synaptic transmission strength. We believe that an optimal fit isn't required or even reasonable, since the datasets are from different individuals and it is not unsurprising that one set of parameters will not perfectly fit data from multiple sources. In addition, the sample size for singly infected cells is large and any comparison between sample and simulated distributions will produce a statistically significant difference if our models do not describe the system perfectly.

We also used our stochastic models to explore the reduction of synaptic strength required to switch between multiply infected and singly infected regimes in the lymph. Stochastic models also suggest the same as our deterministic ones – decreasing the strength of synaptic transmission brings us closer to a majority of singly infected cells. Decreasing the density of cells, or decreasing the probability of successful synaptic transmission can bring us closer to the number of singly infected cells but is not sufficient to switch from one regime to another alone. Another possibility is reducing the number of susceptible cells to synaptic transmission. Although the mechanisms of synaptic transmission are not well understood, target cells may not be susceptible to synaptic transmission due to the complexity required to form a viral synapse. Decreasing the number of susceptible cells can also decrease the strength of synaptic transmission and produce a singly infected regime in the lymph.

Discussion and Conclusion

In experimental studies, a majority of CD4+ T cells in the blood are infected with a single virus while a majority of the CD4+ T cells in the lymph are infected with multiple viruses. We find that this difference can be partially explained by population dynamics rather than any special mechanism. Uninfected, singly infected, and multiply infected cells, as well as free viruses, are released from the lymph to the blood. Since cell-to-cell transmission is less likely in the blood, the infected cells can only transfer viruses through cell-free transmission. This produces singly infected cells in numbers that dwarf the number of multiply infected cells in the blood and masks the number of multiply infected cells.

Through parameter fitting to the observed distributions, we estimate that approximately 2.3 viruses are transmitted per synapse in the lymph. In addition, we estimate that the strength of synaptic transmission at equilibrium is approximately an order of magnitude stronger than free-virus transmission. This relies on the assumption that the number of viruses successfully transferred per synapse follows a normal distribution. There has been suggestion that a few cells successfully transfer hundreds of viruses while most cells do not. However, we do not observe any cells with such a high number of viruses. Cells that receive dozens or hundreds of viruses may not survive to be measured in experimental conditions. Further experimental work will be necessary to measure the number of viruses successfully transferred per synapse and the relative strength of each mode of transmission.

Although population dynamics can explain much of the observed differences in the distribution of singly and multiply infected cells in the blood and lymph, there are more multiply infected cells in the blood than expected in our model than observed. This can be because the sample size is not large enough to detect the small number of highly multiply infected cells in the blood. Another explanation would be bias in the assay against cells that are multiply infected cells. Multiply infected cells may be older or more fragile and less likely to be found by most assays. If the lack of multiply infected cells in the blood is true *in vivo*, then there are several mechanisms that can account for this. We found that higher death rates in highly multiply infected cells, higher death rates in older cells, and spatial release differences can contribute to the lack of multiply infected cells in the blood. We cannot establish the relative contribution of these different factors but show that they are all plausible mechanisms to reduce the number of multiply infected cells in the blood while retaining the number of multiply infected cells in the lymph.

A recent experimental study on infected cells in the lymph node suggests that, contradictory to established data from the spleen, a majority of cells in lymph nodes are singly infected. Our models suggest that this can be due to a reduction in the strength of synaptic transmission, which can be caused by a reduction in the number of susceptible cells. Another possible explanation that we did not explore is faster movement of cells in the lymph nodes versus the spleen. The latter hypothesis is reasonable since the lymph nodes are distributed throughout the body, are closer to the sites of infection, and will

respond to an infection faster than the spleen. Cells in the lymph nodes may be more mobile than those in the spleen.

The work here provides a plausible explanation for the observed difference in multiplicity of infection between the peripheral blood and the spleen. It also provides estimates of key parameter values, including the relative strength of synaptic and free virus transmission. Further experimental work is necessary to measure these parameter values directly and verify the mechanisms predicted in this paper.

SUMMARY AND CONCLUSIONS

This dissertation presents a series of mathematical models that explore different facets of the population dynamics and ecology of HIV within the host. Chapter 1 examines the effect of unintegrated but productive virus on initial viral dynamics. Chapter 2 furthers this by adding an immune response and analyzing the effects of uDNA during the chronic phase of infection. Chapter 3 presents a potential explanation for the observed difference in multiplicity of infection between the lymphoid system and the peripheral blood and provides estimates of key parameter values based on experimental data.

In chapter 1, we develop an ODE model for the effect of unintegrated but productive virus on initial viral dynamics. We derive the expressions for the effect of unintegrated virus on the basic reproductive ratio and determine how factors such as death rate, probability of integration, and infectivity contribute to the basic reproductive ratio of the

virus. We explore the effect of uDNA in the context of free virus transmission and synaptic transmission separately, and combine them to form an overall model for the effect of uDNA on the basic reproductive ratio. We also examine the scenario where multiple unintegrated viruses are required for successful production. Finally, we examine different distributions on the number of viruses transferred per synapse.

We estimate that uDNA can contribute about 20% to the overall basic reproductive ratio of the virus. This figure assumes that a small number of viruses are transferred per synapse, that free and synaptic transmission is equally strong, and the burst size of uDNA infected cells is approximately $1/5^{th}$ that of cells with provirus. Since the reproductive ratio of HIV has been estimated at 4-8, expression from uDNA-only can potentially allow the virus to establish in the host. We also show analytically that if more than one unintegrated virus is required for productive infection, uDNA does not contribute toward the reproductive ratio by free virus infection, but does contribute towards the reproductive ratio by synaptic transmission.

Our model suggests that uDNA can have an effect on initial viral dynamics and can potentially sustain and start infection in the absence of uDNA. This can inform us on the potential efficacy of integrase inhibitors as a prophylactic treatment. In some individuals where the reproductive ratio of HIV is naturally high, integrase inhibitors as a preventative treatment may be ineffective due to production by uDNA-infected cells.

In chapter 2, we employ the same model used in chapter 1, but add an immune response and explore late stage dynamics rather than initial dynamics. We find that that effect of uDNA on steady-state viral loads depends on the strength of viral production from uDNA cells and the strength of the immune response versus these cells. If uDNA-only cells produce high levels of virus and invoke a weak immune response then set point viral loads increase despite an increased immune response. If uDNA-only cells produce relatively low level of virus and invoke a strong immune response, then set point virus levels can actually decrease due to apparent competition. This may be the case *in vivo* since uDNA produce virus at a lower rate yet can present additional viral epitopes that will invoke a stronger immune response versus a particular infected cell. However, no data is available on the strength of the immune response versus productive cells with unintegrated viruses and the case of increased set point viral loads cannot be discounted.

In chapter 3, we use a two-compartment ODE model to try and explain the difference in multiplicity of infection of HIV in the blood and the spleen. A majority of cells in the spleen have more than one virus, while a majority of the cells in the blood have a single virus. However, these systems are connected and we would expect that the cells in both compartments have the same multiplicity of infection. Our models consider multiple infection as well as different modes of transmission. With our ODE model, we show that the observed difference in multiplicity can be explained by a movement of multiply infected cells from the lymph that participate in free virus infection in the blood, producing many singly infected cells.

Our model reproduces the data in the lymph but fails to produce a statistically identical distribution of singly infected cells in the blood. We find that this difference is due to a small number of highly multiply infected cells in our simulations that are not found in the data. Our model predicts that multiply infected cells are rare, but should still be present in the blood in small quantities that was not observed experimentally. This can be due to sampling effects or bias against multiply infected cells in the assay. Under the assumption that there are no multiply infected cells in the peripheral blood *in* vivo, we use a stochastic model to show that an increased death rate of highly multiply infected cells or higher death rate of older cells can all reduce the number of highly multiply infected cells in the blood. Spatial differences in release rates can also favor singly infected cells in the blood if densely packed cells in the interior of the spleen are released at a lower rate. Our models show that population dynamics can explain much of the observed difference in multiplicity but there is likely some other mechanism that reduces the number of highly multiply infected cells in the blood even further.

Our ODE models also suggest that the source of higher multiplicity of infection is synaptic transmission. In the lymph, where cells are dense and relatively immobile, we see a high multiplicity of infection. In the blood, where cells are highly mobile, synaptic transmission is less likely and the multiplicity of infection decreases. We find that by reducing the strength synaptic transmission, we reduce the number of multiply infected cells in the lymph and skew the distribution towards singly infected cells.

Our models also allow us to estimate the average number of viruses transferred per synapse and the strength of synaptic versus free virus transmission. The original ODE model suggests that, on average, 2.3 viruses are transferred per synapse and synaptic transmission is stronger than free virus transmission in the lymph. We use a reduced, steady-state model with less parameters and the results are the same as in our more complicated model. The strength of synaptic transmission and average number of viruses transferred per synapse are fundamental parameters that should be investigated to understand the progression of HIV. If, as our models suggest, synaptic transmission is responsible for more infections than free virus transmission, we should revisit our descriptions of the HIV life cycle. In addition, multiplicity of infection should be explored and understood because it is fundamentally tied to recombination and evolution of the disease within the host.

This dissertation provides a mathematical framework for the study of different aspects of HIV that can serve as a basis for experimentation. We have identified that key parameters in the strength of uDNA infection are the probability of integration and average number of viruses transferred per synapse. These parameters should be directly measured experimentally to quantify the importance of uDNA. We also presented first-order estimates for many parameters based on data from previous experiments. The analysis presented here can serve to guide future experimental work to measure parameters critical for our understanding of the dynamics and progression of HIV.

REFERENCES

[1] Anon. 2013. Global update on HIV Treatment 2013: Results, Impact and Opportunities. WHO report, in partnership with UNICEF and UNAIDS. Geneva: WHO.

[2] Merson, MH. 2006. The HIV-AIDS Pandemic at 25 – The Global Response. N Engl J Med 2006; 354:2414-2417.

[3] Anon. 2001. Morbidity and Mortality Weekly Report: HIV and AIDS – United States, 1981-2000. CDC.

[4] Mathias, A. A. et al. 2007. Bioequivalence of efavirenz/emtricitabine/tenofovir disoproxil fumarate single-tablet regimen. J. Acquir. Immune Defic. Syndr. 46, 167–173

[5] Cohen, M S. et al. 2011. Prevention of HIV-1 Infection with Early Antiretroviral Therapy. The New England Journal of Medicine 365:493-505.

[6] Mugavero, MJ, et al. 2013. The state of engagement in HIV care in the United States: from cascade to continuum to control. Clinical infectious diseases 57.8: 1164-1171.

[7] Frenkel, L. M. et al. 2003. Multiple viral genetic analyses detect low-level human immunodeficiency virus type 1 replication during effective highly active antiretroviral therapy. J. Virol. 77,5721–5730

[8] Ho, DD, et al. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. Jan 12;373(6510):123-6.

[9] Wei, X, et al. 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature. Jan 12;373(6510):117-22.

[10] Hellerstein, M. 1999. Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. Nat Med. Jan;5(1):83-9.

[11] Perelson, AS. 1996. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. Science.

[12] Ladell, K, et al. 2008. Central Memory CD8+ T Cells Appear to Have a Shorter Lifespan and Reduced Abundance as a Function of HIV Disease Progression. The Journal of ImmunologyJune 15, vol. 180 no. 12 7907-7918.

[13] Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Lori F, Flexner C, Quinn TC, Chaisson RE, Rosenberg E, Walker B, Gange S,

Gallant J, Siliciano RF. 1999. Latent infection of CD4+ T-cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy.
Nat Med, 5:512-517.

[14] Hu, WS. Temin, HM. 1990. Retroviral recombination and reverse transcription. Science 30 November.

[15] Engelman, A., Cherepanov, P. 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. Nature Reviews Microbiology 10, 279-290.

[16]. Feng Y, Broder C C, Kennedy P E, Berger E A. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.Science.272:872–877.

[17] Bleul CC, Wu L, Hoxie JA, Springer TA, Mackay CR. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. 1997. Proceedings of the National Academy of Sciences of the United States of America ;94(5):1925-1930.

[18] Bukrinsky, Michael I., et al. 1992. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. Proceedings of the National Academy of Sciences 89.14 : 6580-6584.

[19] Levy, David N, et al. 2004. Dynamics of HIV-1 recombination in its natural target cells. PNAS 101 (12) 4204-4209.

[20] Tersmette, M., Lange, J.M., de Goede, R.E., de Wolf, F.,et al. 1989.
Association between biological properties of human immunodifciency virus variants and risk for AIDS and AIDS mortality. Lancet 1, 983-5.

[21] Moore, J.P., Trkola, A., Dragic, T. 1997. Co-receptors for HIV-1 entry. Current Opinion Immunology 9, 551-62.

[22] Chun TW, Carruth L, Finzi D, Shen X, DiGiuseppe JA, Taylor H, Hermankova M, Chadwick K, Margolick J, Quinn TC, Kuo YH, Brookmeyer R, Zeiger MA, Barditch-Crovo P, Siliciano RF. 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection.
Nature, 387:183-188.

[23] Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, Lloyd AL, Nowak MA, Fauci AS. 1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy.
Proc Natl Acad Sci U S A, 94:13193-13197.

[24] Chun TW, Finzi D, Margolick J, Chadwick K, Schwartz D, Siliciano RF. 1995. In vivo fate of HIV-1-infected T-cells: quantitative analysis of the transition to stable latency.

Nat Med, 1:1284-1290.

[25] Katlama C, Deeks SG, Autran B, Martinez-Picado J, van Lunzen J, Rouzioux C, Miller M, Vella S, Schmitz JE, Ahlers J, Richman DD, Sekaly RP. 2013. Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. Lancet, 381:2109-2117

[26] Dang, Q., J. Chen, D. Unutmaz, J. M. Coffin, V. K. Pathak, D. Powell, V. K. KewalRamani, F. Maldarelli, and W.-S. Hu. 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc. Natl. Acad. Sci. USA 101:632-637.

[27] Dang, Q., J. Chen, D. Unutmaz, J. M. Coffin, V. K. Pathak, D. Powell, V. K. KewalRamani, F. Maldarelli, and W.-S. Hu. 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc. Natl. Acad. Sci. USA 101:632-637.

[28] Gratton, S., Cheynier, R., Dumaurier, M. J., Oksenhendler, E. & Wain-Hobson, S. 2002. Highly restricted spread of HIV-1 and multiply infected cells within splenic germinal centers. Proc. Natl Acad. Sci. USA 97, 14566–14571.

[29] Jung, A., Maier, R., Vartanian, J.-P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S. & Meyerhans, A. 2002. Recombination: Multiply infected spleen cells in HIV patients Nature 418 , 144.

[30] Levy, D. N., Aldrovandi, G. M., Kutsch, O. & Shaw, G. M. 2004. Dynamics of HIV-1 recombination in its natural target cells. Proc. Natl. Acad. Sci. USA 101 ,4204-4209, and correction 102, 1808.

[31] Dang, Q., Chen, J., Unutmaz, D., Coffin, J. M., Pathak, V. K., Powell, D., Kewa Ramani, V. N., Malderalli, F. & Hu, W.-S. 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc. Natl. Acad. Sci. USA 101 , 632-637.

[32] Chen, J., Dang, Q., Unutmaz, D., Pathak, V. K., Maldarelli, F., Powell, D. & Hu, W.-S. 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.

[33] Josefsson, et al. 2011. Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV. Proc Natl Acad Sci U S A. 2011 Jul 5;108(27):11199-204

[34] Josefsson, L. et al. 2013. 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 Pathogens.

[35] Chen, B. K., Gandhi, R. T. & Baltimore, D. 1996. CD4 down-modulation during infection of human T cells with human immunodeficiency virus type 1 involves independent activities of vpu, env, and nef. J. Virol. 70, 6044–6053

[36] Dixit, N., Perelson, A. 2005. HIV dynamics with multiple infections of target cells. PNAS. 8198–8203, doi: 10.1073/pnas.0407498102.

[37] Bangham, C. R. 2003. The immune control and cell-to-cell spread of human T-lymphotropic virus type 1. J. Gen. Virol. 84:3177-3189.

[38] Jolly, C., and Q. J. Sattentau. 2004. Retroviral spread by induction of virological synapses. Traffic 5:643-650.

[39] Feldmann J, Schwartz O. 2010. HIV-1 virological synapse: live imaging of transmission. Viruses, 2:1666-1680

[40] Sattentau QJ. 2010. Cell-to-cell spread of retroviruses. Viruses, 2:1306-1321.

[41] Dale BM, Alvarez RA, Chen BK. 2013. Mechanisms of enhanced HIV spread through T-cell virological synapses. Immunol Rev, 251:113-124.

[42] Gelderblom, H. et al. 2008. Viral Complementation allows HIV-1 replication without integration. Retrovirology 5:60.

[43] Trinite B, Ohlson EC, Voznesensky I, Rana SP, Chan CN, et al. 2013. An HIV-1 Replication Pathway Utilizing Reverse Transcription Products That Fail To Integrate. J Virol 87: 12701-12720.

[44] Sloan, R., Wainberg, M. 2011. The role of unintegrated DNA in HIV infection. Retrovirology.

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

[46]. Perelson AS, Ribeiro RM. 2002. Modelling viral and immune system dynamics. Nature Rev Immunol 2: 28-36. 3.

[47] Perelson, AS. 2013. Modeling the within-host dynamics of HIV infection. BMC Biol 11: 96. 4.

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

[49] Perelson, A, et al. 2013. Modeling the within-host dynamics of HIV infection. BMC Biology , 11:96.

[50] Levin S.A., Grenfell, B., Hastings, A., Perelson, A.S. 1997. Mathematical and computational challenges in population biology and ecosystems science. Science 275, 334-343. 1997.

[51] Lotka, A.J. 1956. Elements of mathematical biology. Dover, New York. Dover Publishing.

[52] Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., et al. 1997. Decay characteristics of HIV-1 infected compartments during combination therapy. Nature 387, 188-91.

[53] Bonhoeffer, S., May, R.M., Shaw, G.M., Nowak, M.A. 1997. Virus dynamics and drug therapy. PNAS 94, 6971-6976.

[54] Bonhoeffer, S., Nowak, M.A. 1997. Pre-existence and emergence of drug resistance in HIV-1 infection. ProcR. Soc Lond B Biology 264, 631-7.

[55] Ribeiro, R.M., Bonhoeffer, S. 2000. Production of resistant HIV mutants during antiretroviral therapy. PNAS 97, 7681-6.

[56] Anderson, R.M., May, R.M. 1991. Infectious diseases of humans. Oxford, England. Oxford University Press.

[57]. Ribeiro RM, Qin L, Chavez LL, Li D, Self SG, et al. 2010. Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. J Virol 84: 6096-6102.

[58] Sloan, R., Wainberg, M. 2011. The role of unintegrated DNA in HIV infection. Retrovirology.

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

[60] Komarova, NL, Anghelina, D. Voznesensky, I, Trinite, B. Levy, DN, et al. 2013. Relative contributions of free-virus and synaptic transmission to the spread of HIV-1 through target cell populations. Biology Letters 9: 20121049 .

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

[62] Komarova, NL, Wodarz D. 2013. Virus dynamics in the presence of synaptic transmission. Math Biosci 242: 161-171.

[63]. Llibre JM, Buzón MJ, Massanella M, Esteve A, Dahl V, et al. 2012. Treatment intensification with raltegravir in subjects with sustained HIV-1 viraemia suppression: a randomized 48-week study. Antiviral therapy 17: 355.

[64] Buzón MJ, Massanella M, Llibre JM, Esteve A, Dahl V, et al. 2010. HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed

subjects. Nature Medicine 16: 460-465.

[65] Bushman, F. Craigie, R. 2012. HIV DNA Integration. Cold Spring Harb Perspect Med.; 2(7): a006890.

[66] Ridge, J.P., Di Rosa, F., Matzinger, P. 1998. A conditioned dendritic cell can be a temporal bridge between a CD4+ T-helper and a T-killer cell. Nature 393, 474-8.

[67] Schoenberger, S.P., Toes, R.E., et al. 1998. T-cell help for cytotoxic T lymphocytes is mediated by CD40-CD40L interactions. Nature 393, 480-3.

[68] Kagi, D., Hengartner, H. 1996. Different Roles for Cytotoxic T-Cells in the Control of Infections With Cytopathic Versus Noncytopathic Viruses. Current opinion in Immunology 8, 472-477.

[69] Kagi, D., Ledermann, B., Burki, K., Zinkernagel, R.M., Hengartner, H. 1995. Lymphocyte Mediated Cytotoxicity in-vitro and in-vivo – Mechanisms and Significance. Immunological Reviews 146, 95-155.

[70] Kagi, D., et al. 1996. Molecular Mechanisms of Lymphocyte-Mediated Cytotoxcity and Their Role in Immunological Protection and Pathogenesis in vivo. Annual Review of Immunology 14, 207-232.

[71] Kagi, D. Seiler, P. et al. 1995. The Roles of Perferin-Dependent and Fas-Dependent Cytotoxicity in Protection and Against Cytopathic and Noncytopathic Viruses. European Journal of Immunology 25, 2356-3262.

[72] de Boer, R.J., Boerlijst, M.C. 1994. Diversity and virulence thresholds in AIDS. PNAS 91 544-8.

[73] de Boer, R.J., Perelson, A.S. 1994. T-cell repertoires and competitive exclusion. J. Theor Biol 169, 375-90.

[74] de Boer, R.J. Perelson, A.S. 1995. Towards a general function describing T cell proliferation. J Theor Biol 175, 567-76.

[75] De Boer, R.J., Perelson, A.S. 1998. Target cell limited and immune control models of HIV infection: a comparison. J Theor Biol 190, 201-14.

[76] Callaway, Duncan S., and Alan S. Perelson. 2002. HIV-1 infection and low steady state viral loads. Bulletin of mathematical biology 64.1 : 29-64.

[77] Dixit NM, Perelson AS. 2004. Multiplicity of human immunodeficiency virus infections in lymphoid tissue. J Virol, 78:8942-8945.

[78] Martin, N. et al. 2010. Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. J. Virol. 84, 3516–3527

[79] Funk, GA. 2001. Quantification of in vivo replicative capacity of HIV-1 in different compartments of infected cells. Journal of Acquired Immune Deficiency Syndrome. Apr 15;26(5):397-404.

[80] Westermann J, Pabst R. 1992. Distribution of lymphocyte subsets and natural killer cells in the human body. Clin Investig 70: 539–544

[81] Haase, a T. 1999. Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphoid tissues. Annual review of immunology, 17, 625–56. doi:10.1146/annurev.immunol.17.1.625.

[82] Sourisseau, M., Sol-Foulon, N., Porrot, F., Blanchet, F. & Schwartz, O. 2007. Inefficient human immunodeficiency virus replication in mobile lymphocytes. J. Virol. 81, 1000–1012

[83] Hubner, W. et al. 2009. Quantitative 3D Video Microscopy of HIV Transfer Across T Cell Virological Synapses. Science.

[84] Hankins, C. 2013. Overview of the current state of the epidemic. Current HIV/AIDS Reports 10.2 113-123.

[85] Holt, RD, Lawton, JH. Apparent Competition and Enemy-Free Space in Insect Host-Parasitoid Communities. 1993. The American Naturalist, Vol. 142, No. 4 pp. 623-645

Appendix A
Mathematical details of the models discussed in Chapter 1

Introduction

This document presents the model and derivation of the various R_0 presented in the paper. First, we introduce the overall model with both synaptic and free-virus transmission. We calculate R_0 for the sub-model which only includes free-virus transmission. Then, we calculate R_0 for the model with just synaptic transmission. We provide numerical results to support the combined R_0 for the complete model. Finally, we explore different formulations for the rate of synaptic transmission.

Complete Model with Synaptic and Free Virus Transmission

$$\frac{dy_{0,0}}{dt} = \lambda - dy_{0,0} - \beta y_{0,0} v - y_{0,0} \sum_{k=0}^{N} \sum_{l=0}^{M} y_{kl} \sum_{r=0}^{N} \sum_{s=0}^{M} \gamma_{rs}^{kl}$$

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

$$\frac{dv}{dt} = \sum_{i=0}^{N} \sum_{j=0}^{M} k_{ij} y_{ij} - uv$$

Notation:

y_{ij} A cell with Unintegrated (i), Integrated (j) viruses
N, maximum number of unintegrated viruses per cell
M, maximum number of integrated viruses per cell
k,r: dummy variables used as counters for integrated viruses
l,s: dummy variables used as counters for unintegrated viruses

Parameters:

λ : Birth rate of uninfected CD4+ T-cells
d : Death rate of uninfected CD4+ T-cells.
β : Free Virus Infectivity - effectiveness of the interaction between virus and a CD4+ T cell
p : Probability of integration.
$k_{i,j}$: Virus replication parameter - rate of free virus production by an infected cell

u : Clearance rate of free virus
$a_{i,j}$: Clearance rate of infected cells

$a_{i,j}$ can be defined as:

$$a_{ij} = \begin{cases} d & \text{if } j = 0, i < c \\ a_u & \text{if } j = 0, i \geq c \\ a_i & \text{if } j > 0 \end{cases}$$

$k_{i,j}$ can be defined as:

$$k_{ij} = \begin{cases} 0 & \text{if } j = 0, i < c \\ k_u & \text{if } j = 0, i \geq c \\ k_i & \text{if } j > 0 \end{cases}$$

c : the threshold number of uDNA viruses necessary for uDNA-only cells to begin production.

γ_{rs}^{kl} : the probability that a cell with k integrated, l unintegrated viruses transfers r integrated and s unintegrated viruses. We set $\gamma_{00}^{kl} = 0$ and $\gamma_{rs}^{00} = 0$

$\sum_{k=1}^{N} \sum_{l=1}^{M} y_{kl} (\sum_{r=1}^{i} \sum_{s=1}^{j} \gamma_{rs}^{kl} y_{i-r,j-s})$ Represents cells transmitting to other cells to create y_{ij}. The second set of summations only reach i and j because a cell can only transmit up to i and j viruses to an empty cell to get y_{ij}. The y_{kl} term represents the transmitting (source) cell. Any productively infected cell can be a source cell. $y_{i-r,j-s}$ represents the target cell that accepts viruses.

$\sum_{k=1}^{N} \sum_{l=1}^{M} y_{kl} (y_{ij} \sum_{r=1}^{N-i} \sum_{s=1}^{M-j} \gamma_{rs}^{kl})$:Represents cells transmitting to y_{ij}

Note that γ_{rs}^{00}, γ_{00}^{kl}, and k_{00} are all zero. These are synaptic transmission rates and viral production, respectively.

R_0 Calculation: Free-Virus Transfer Alone

$$\frac{dy_{0,0}}{dt} = \lambda - dy_{0,0} - \beta y_{0,0} v$$

$$\frac{dy_{i,j}}{dt} = (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j}$$

$$\frac{dv}{dt} = \sum_{i=1}^{N} \sum_{j=1}^{M} k_{ij} y_{ij} - uv$$

The virus-free equliibrium of the above system is $y_{00} = \lambda/d, y_{ij} = 0, v = 0$. With standard methods, we can calculate the Jacobian of this system at this steady state to arrive at R_0. Below is the Jacobian of the system with the assumption that there are two distinct death rates and two distinct virus production rates. Cells infected with only unintegrated viruses produce virus and die at a lower rate then the cells infected with provirus.

$$J_{NM+1\times NM+1} = \begin{pmatrix} -d & 0 & 0 & ... & 0 & 0 & ... & 0 & -\frac{\beta\lambda}{d} \\ 0 & -a_{\text{un}} & 0 & ... & 0 & 0 & ... & 0 & -\frac{(-1+p)\beta\lambda}{d} \\ 0 & 0 & -a_{\text{un}} & ... & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & -a_{\text{in}} & 0 & ... & 0 & \frac{p\beta\lambda}{d} \\ 0 & 0 & 0 & ... & 0 & -a_{\text{in}} & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 0 & ... & -a_{\text{in}} & 0 \\ 0 & k_{\text{un}} & k_{\text{un}} & ... & k_{\text{in}} & k_{\text{in}} & ... & k_{\text{in}} & -u \end{pmatrix}$$

All but 3 of the eigenvalues for this system are $-d, -a_i$ and $-a_u$. The remaining eigenvalues are the roots of the 3rd degree polynomial:

$$p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

Where

$$a_3 = d$$
$$a_2 = du + da_{\text{in}} + da_{\text{un}}$$
$$a_1 = dua_{\text{un}} + da_{\text{in}}\left(u + a_{\text{un}}\right) - \beta\lambda\left(pk_{\text{in}} - (p-1)k_{\text{un}}\right)$$
$$a_0 = p\beta\lambda a_{\text{un}}k_{\text{in}} - a_{\text{in}}\left(dua_{\text{un}} + (-1+p)\beta\lambda k_{\text{un}}\right)$$

These eigenvalues are nonnegative when they satisfy the Routh-Hurwitz criterion, which occurs whenever the following inequality is satisfied:

$$R_0 \text{ Free Virus Only, No Threshold } : \frac{p\beta\lambda k_i}{ua_id} + \frac{(1-p)\beta\lambda k_u}{ua_ud} > 1$$

R_0 Calculation: Free-Virus Transfer with threshold for active uDNA

If a cell without an integrated provirus requires more than one unintegrated virus to begin production then the system of ODE remains the same, except the virus production rate (k_u) is zero for cells without integrated viruses and

below a threshold number of unintegrated viruses. In addition, these cells have a clearance rate equal to that of uninfected cells.

The Jacobian for this system becomes:

$$J_{NM+1\times NM+1} = \begin{pmatrix} -d & 0 & ... & 0 & ... & 0 & 0 & ... & 0 & -\frac{\beta\lambda}{d} \\ 0 & -d & ... & 0 & ... & 0 & 0 & ... & 0 & -\frac{(-1+p)\beta\lambda}{d} \\ 0 & 0 & \ddots & 0 & ... & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & ... & -a_{\text{un}} & ... & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & ... & 0 & \ddots & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & ... & 0 & ... & -a_{\text{in}} & 0 & ... & 0 & \frac{p\beta\lambda}{d} \\ 0 & 0 & ... & 0 & ... & 0 & -a_{\text{in}} & ... & 0 & 0 \\ 0 & 0 & ... & 0 & ... & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & ... & 0 & ... & 0 & 0 & ... & -a_{\text{in}} & 0 \\ 0 & 0 & ... & k_{\text{un}} & ... & k_{\text{in}} & k_{\text{in}} & ... & k_{\text{in}} & -u \end{pmatrix}$$

With the requirement that more than one uDNA virus is necessary to begin virus production, all but 2 of the eigenvalues are $-d, -a_i$ or $-a_u$. The last remaining eigenvalues are:

$$\frac{1}{2}\left(-u - a_{\text{in}} \pm \frac{\sqrt{du^2 - 2dua_{\text{in}} + da_{\text{in}}^2 + 4p\beta\lambda k_{\text{in}}}}{\sqrt{d}}\right)$$

which is non-negative whenever:

$$R_0 \text{ Free Virus Only, With Threshold } = \frac{p\beta\lambda k_i}{ua_i d} > 1$$

R_0 Calculation: Synaptic Transfer Alone

$$\frac{dy_{0,0}}{dt} = \lambda - dy_{0,0} - y_{0,0}\sum_{k=0}^{N}\sum_{l=0}^{M} y_{kl}\sum_{r=0}^{N}\sum_{s=0}^{M}\gamma_{rs}^{kl}$$

$$\frac{dy_{i,j}}{dt} = -a_{i,j}y_{i,j} + \sum_{k=0}^{N}\sum_{l=0}^{M} y_{kl}\left(\sum_{r=0}^{i}\sum_{s=0}^{j}\gamma_{rs}^{kl}y_{i-r,j-s} - y_{ij}\sum_{r=0}^{N-i}\sum_{s=0}^{M-j}\gamma_{rs}^{kl}\right)$$

The virus-free equliibrium of the above system is $y_{00} = \lambda/d, y_{ij} = 0$. With standard methods, we can calculate the Jacobian of this system at this steady state to arrive at R_0. Below is the Jacobian of the system with the assumption

that there are two death rates: death rates for cells infected with uDNA only and a separate death rate for iDNA infected cells. Note that we make no assumption about the values of γ.
The Jacobian for the system with synaptic transfer alone is an NM by NM matrix:

$$J = \begin{pmatrix} -d & -\frac{\lambda}{d}\sum\frac{\lambda}{d}\gamma_{ij}^{01} & \sum\frac{\lambda}{d}\gamma_{ij}^{02} & \cdots & \sum\frac{\lambda}{d}\gamma_{ij}^{10} & \sum\frac{\lambda}{d}\gamma_{ij}^{11} & \cdots & \sum\frac{\lambda}{d}\gamma_{ij}^{NM} \\ 0 & -a_u+\frac{\lambda}{d}\gamma_{01}^{01} & \frac{\lambda}{d}\gamma_{01}^{02} & \cdots & \frac{\lambda}{d}\gamma_{01}^{10} & \frac{\lambda}{d}\gamma_{01}^{11} & \cdots & \frac{\lambda}{d}\gamma_{01}^{NM} \\ 0 & \frac{\lambda}{d}\gamma_{02}^{01} & -a_u+\frac{\lambda}{d}\gamma_{02}^{02} & \cdots & \frac{\lambda}{d}\gamma_{02}^{10} & \frac{\lambda}{d}\gamma_{02}^{11} & \cdots & \frac{\lambda}{d}\gamma_{02}^{NM} \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \frac{\lambda}{d}\gamma_{10}^{01} & \frac{\lambda}{d}\gamma_{10}^{02} & \cdots & -a_i+\frac{\lambda}{d}\gamma_{10}^{10} & \frac{\lambda}{d}\gamma_{10}^{11} & \cdots & \frac{\lambda}{d}\gamma_{10}^{NM} \\ 0 & \frac{\lambda}{d}\gamma_{11}^{01} & \frac{\lambda}{d}\gamma_{11}^{02} & \cdots & \frac{\lambda}{d}\gamma_{10}^{10} & -a_i+\frac{\lambda}{d}\gamma_{11}^{11} & \cdots & \frac{\lambda}{d}\gamma_{11}^{NM} \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \frac{\lambda}{d}\gamma_{NM}^{01} & \frac{\lambda}{d}\gamma_{NM}^{02} & \cdots & \frac{\lambda}{d}\gamma_{NM}^{10} & \frac{\lambda}{d}\gamma_{NM}^{11} & \cdots & -a_i+\frac{\lambda}{d}\gamma_{NM}^{NM} \end{pmatrix}$$

where , $\sum = \sum_{i=0}^{N}\sum_{j=0}^{M}$:

The first eigenvalue of the system is -d. To determine the remaining eigenvalues we look at the remaining NM-1 by NM-1 sub matrix:

$$\hat{J} = A_u + \begin{pmatrix} \frac{\lambda}{d}\gamma_{01}^{01} & \frac{\lambda}{d}\gamma_{01}^{02} & \cdots & \frac{\lambda}{d}\gamma_{01}^{10} & \frac{\lambda}{d}\gamma_{01}^{11} & \cdots & \frac{\lambda}{d}\gamma_{01}^{NM} \\ \frac{\lambda}{d}\gamma_{02}^{01} & \frac{\lambda}{d}\gamma_{02}^{02} & \cdots & \frac{\lambda}{d}\gamma_{02}^{10} & \frac{\lambda}{d}\gamma_{02}^{11} & \cdots & \frac{\lambda}{d}\gamma_{02}^{NM} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\lambda}{d}\gamma_{10}^{01} & \frac{\lambda}{d}\gamma_{10}^{02} & \cdots & q+\frac{\lambda}{d}\gamma_{10}^{10} & \frac{\lambda}{d}\gamma_{10}^{11} & \cdots & \frac{\lambda}{d}\gamma_{10}^{NM} \\ \frac{\lambda}{d}\gamma_{11}^{01} & \frac{\lambda}{d}\gamma_{11}^{02} & \cdots & \frac{\lambda}{d}\gamma_{10}^{10} & q+\frac{\lambda}{d}\gamma_{11}^{11} & \cdots & \frac{\lambda}{d}\gamma_{11}^{NM} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\lambda}{d}\gamma_{NM}^{01} & \frac{\lambda}{d}\gamma_{NM}^{02} & \cdots & \frac{\lambda}{d}\gamma_{NM}^{10} & \frac{\lambda}{d}\gamma_{NM}^{11} & \cdots & q+\frac{\lambda}{d}\gamma_{NM}^{NM} \end{pmatrix}$$

where $q = a_u - a_i$ and $A_u = a_u I$.

This matrix has (NM-1)(NM-1)+3 different variables and to simplify it we assume that the transfer rate γ_{rs} is constant for all iDNA-infected cells, but for uDNA-infected cells the rate is $\eta\gamma_{rs}$, where $\eta < 1$

Under these assumptions, we can replace the above Jacobian by:

$$\hat{J} = A_u + \begin{pmatrix} \frac{\lambda}{d}\gamma_{01} & \frac{\lambda}{d}\gamma_{01} & \cdots & \frac{\lambda}{d}\gamma_{01} & \frac{\lambda}{d}\gamma_{01} & \cdots & \frac{\lambda}{d}\gamma_{01} \\ \frac{\lambda}{d}\gamma_{02} & \frac{\lambda}{d}\gamma_{02} & \cdots & \frac{\lambda}{d}\gamma_{02} & \frac{\lambda}{d}\gamma_{02} & \cdots & \frac{\lambda}{d}\gamma_{02} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\lambda}{d}\gamma_{10} & \frac{\lambda}{d}\gamma_{10} & \cdots & q+\frac{\lambda}{d}\gamma_{10} & \frac{\lambda}{d}\gamma_{10} & \cdots & \frac{\lambda}{d}\gamma_{10} \\ \frac{\lambda}{d}\gamma_{11} & \frac{\lambda}{d}\gamma_{11} & \cdots & \frac{\lambda}{d}\gamma_{10} & q+\frac{\lambda}{d}\gamma_{11} & \cdots & \frac{\lambda}{d}\gamma_{11} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\lambda}{d}\gamma_{NM} & \frac{\lambda}{d}\gamma_{NM} & \cdots & \frac{\lambda}{d}\gamma_{NM} & \frac{\lambda}{d}\gamma_{NM} & \cdots & q+\frac{\lambda}{d}\gamma_{NM} \end{pmatrix}$$

Note that if we require a threshold number (c) of uDNA viruses to start production, the leftmost c columns become zero.
All but 2 eigenvalues of this matrix is q or 0. The remaining two eigenvalues are:

$$\frac{1}{2} * (\frac{\lambda}{d}\beta_{su} + \frac{\lambda}{d}\beta_{si} + q \pm \sqrt{-4 * \frac{\lambda}{d}\beta_{su}q + (\frac{\lambda}{d}\beta_{su} + \frac{\lambda}{d}\beta_{si} + q)^2}) - a_u$$

Where $\beta_{su} = \sum_{r=1}^{N} \eta\gamma_{r,0}$ (zero integrated, r unintegrated, r $\geq$ c) and $\beta_{si} = \sum_{r=0}^{N}\sum_{s=1}^{N}\gamma_{r,s}$. This accounts for the reduction in transmission from uDNA-only cells.

These eigenvalues are exactly 0 when:

$$\frac{\lambda}{d}a_u\beta_i + \frac{\lambda}{d}a_i\beta_u = a_i a_u$$

And the new R_0 is now:

$$R_0 \text{ Synaptic Only : } \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

R_0 Calculation: Combined Cell Free and Synaptic Transmission

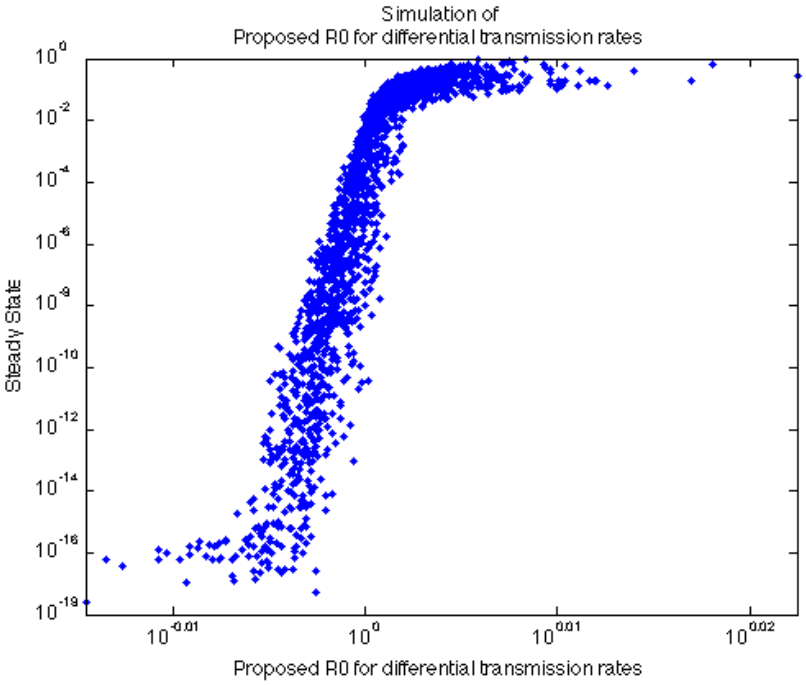
Although the analytical solution of R_0 of the combined system of ODEs is unknown, numerical simulations (See figure) show that the critical value is:

$$R_0 \text{ Complete Model, No Threshold : } \frac{p\beta\lambda k_i}{ua_i d} + \frac{(1-p)\beta\lambda k_u}{ua_u d} + \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

If more than one integrated viruses are required to begin infection, this becomes:

$$R_0 \text{ Complete Model, Threshold : } \frac{p\beta\lambda k_i}{ua_i d} + \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

Figure 1: Simulation of the combined system show R_0 approximately as predicted analytically - all parameters randomized. As R_0 approaches 1 from below, the system requires more time to approach equilibrium. There is a sharp transition between zero steady states and non-zero steady states when the proposed R_0 is above 1.



Expanding γ_{rs}

If we assume that viruses integrate independently, then for any number of viruses transferred, $h = r+s$, the chance of any one of these viruses becoming integrated is p and the chance of any single one of these viruses becoming unintegrated is (1-p). If we assume that these are independent events, then the resulting distribution of these h viruses follows a binomial distribution. In addition, each γ is can be decomposed to the probability to transfer h viruses and the probability that some portion of the h viruses become integrated.

Then we can define a set of transfer rates for any given number of cells:

$$\hat{\gamma}^h \text{ for } h = 1....\max(N, M)$$

Note that for transfer from cells with only unintegrated viruses, the transfer rate is reduced by η. This defines the transfer rate of h viruses that become r unintegrated and s integrated viruses, where the transfer rates are:

$$\gamma_{r,0} = \hat{\gamma}^r (1-p)^r$$

$$\gamma_{r,s} = \hat{\gamma}^{r+s} \binom{r+s}{r} (1-p)^r p^s$$

The first represents r viruses being transferred and all r becoming unintegrated. The second represents a total of r+s viruses transferred through the synapse and s viruses becoming integrated while r viruses are unintegrated.

Then:

$$\beta_{su} = \eta \sum_{r=c}^{N} \hat{\gamma}^r (1-p)^r$$

$$\beta_{si} = \sum_{r=0}^{N} \sum_{s=1}^{M} \hat{\gamma}^{r+s} \binom{r+s}{r} (1-p)^r p^s$$

Which can be simplified to:

$$\beta_{su} = \eta \sum_{r=c}^{N} \hat{\gamma}^r (1-p)^r$$

$$\beta_{si} = \sum_{r=1}^{N} \hat{\gamma}^r (1-(1-p)^r)$$

Therefore R_0 of the model with synaptic transfer only is:

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

And R_0 of the complete model is:

$$R_0 = \frac{p\beta\lambda k_i}{u a_i d} + \frac{\lambda}{a_i d} \hat{\gamma}^r \sum_{r=1}^{N} (1-(1-p)^r) + \frac{\eta\lambda}{a_u d} \hat{\gamma}^r \sum_{r=c}^{N} (1-p)^r$$

Finally, we explore different distributions for $\hat{\gamma}^h$, for h=1....N. It is unclear how many viruses are transferred per synapse, so we explore different distributions for the number of viruses transferred.

Uniform: γ^h is the same for all h

$$\text{Normal: } \gamma^h \sim \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(h-\hat{h})^2}{2\sigma^2}}$$

$$\text{Poisson: } \gamma^h \sim \frac{\hat{h}^h \exp^{-\hat{h}}}{h!}$$

For example, for the poisson distribution, the final value for R_0 becomes:

$$R_0 = \frac{p\beta\lambda k_i}{ua_i d} + \frac{\lambda}{a_i d}\frac{\hat{h}^r \exp^{-\hat{r}}}{r!}\sum_{r=1}^{N}(1-(1-p)^r) + \frac{\eta\lambda}{a_u d}\frac{\hat{h}^r \exp^{-\hat{h}}}{r!}\sum_{r=c}^{N}(1-p)^r$$